

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:38:47 ; Search time 8472.11 Seconds
(without alignments)
12117.322 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggccggcct.....tgtcgaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_htc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query			Description
			%	Match	Length	
1	960.2	35.6	2802	9	AY406491	AY406491 Homo sapi
2	950.4	35.2	2791	9	AY406493	AY406493 Mus muscu
3	923.4	34.2	3790	3	AK031655	AK031655 Mus muscu
4	874	32.4	1852	3	CR598115	CR598115 full-leng
5	871.4	32.3	3866	3	AK018177	AK018177 Mus muscu
6	814	30.2	2802	9	AY406492	AY406492 Pan trogl
7	810.4	30.0	2532	9	AY411747	AY411747 Homo sapi
8	780.4	28.9	2532	9	AY411749	AY411749 Mus muscu

c	9	768.6	28.5	1034	4	BI758231	BI758231	603029876
	10	736.6	27.3	1532	3	BC033727	BC033727	Homo sapi
	11	735.2	27.3	788	1	AI951556	AI951556	wv36f04.x
	12	721.6	26.8	796	5	BX348193	BX348193	BX348193
	13	713.4	26.5	818	4	BI818609	BI818609	603033362
	14	678.2	25.1	2532	9	AY411748	AY411748	Pan trogl
	15	672.6	24.9	934	2	BF311804	BF311804	601897316
	16	665	24.7	859	2	BF311896	BF311896	601897733
	17	647.4	24.0	853	5	BX364574	BX364574	BX364574
	18	621	23.0	900	5	BX345406	BX345406	BX345406
	19	618.8	22.9	756	5	BU612387	BU612387	UI-M-EW0-
	20	610.2	22.6	2775	9	AY401471	AY401471	Mus muscu
	21	603.2	22.4	874	5	BQ689148	BQ689148	AGENCOURT
	22	600.6	22.3	889	5	BQ691915	BQ691915	AGENCOURT
	23	600.2	22.3	977	5	BX345407	BX345407	BX345407
	24	599	22.2	2775	9	AY401469	AY401469	Homo sapi
	25	595.2	22.1	601	1	AL516580	AL516580	AL516580
	26	589.4	21.9	604	2	BE314370	BE314370	601147261
	27	586.6	21.8	1072	5	BX422753	BX422753	BX422753
	28	566.6	21.0	678	6	CA749784	CA749784	UI-M-FD0-
	29	561	20.8	1175	2	BF530640	BF530640	602071931
	30	555.6	20.6	788	6	CA317532	CA317532	UI-M-FW0-
	31	555	20.6	572	7	CR554569	CR554569	DKFZp459I
	32	491.2	18.2	499	7	CR747398	CR747398	CR747398
	33	486.2	18.0	548	5	BX452510	BX452510	BX452510
	34	471.2	17.5	540	4	BI467799	BI467799	389583 MA
	35	470.8	17.5	2507	9	AY401470	AY401470	Pan trogl
	36	468.6	17.4	824	4	BI737024	BI737024	603360874
	37	438.2	16.2	675	6	CA315487	CA315487	UI-M-FW0-
	38	433	16.1	471	5	BX282095	BX282095	BX282095
	39	428.4	15.9	1147	5	BU840446	BU840446	AGENCOURT
	40	424	15.7	843	4	BG913440	BG913440	602811321
	41	417.8	15.5	856	7	CN164143	CN164143	994266 MA
	42	411.8	15.3	749	7	CF735417	CF735417	UI-M-HB0-
	43	410.6	15.2	1111	4	BG298307	BG298307	602397080
	44	409.8	15.2	751	7	CF735550	CF735550	UI-M-HB0-
	45	408.8	15.2	460	2	BF443156	BF443156	BF443156 260542 MA

ALIGNMENTS

RESULT 1

AY406491

LOCUS AY406491 2802 bp DNA linear GSS 15-DEC-2003
 DEFINITION Homo sapiens UNCSC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY406491
 VERSION AY406491.1 GI:39762465
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2802)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2802)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
 source Location/Qualifiers
 1. .2802
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1. .>2802
 /gene="UNC5C"
 /locus_tag="HCM2575"
ORIGIN
 Query Match 35.6%; Score 960.2; DB 9; Length 2802;
 Best Local Similarity 61.3%; Pred. No. 1.9e-194;
 Matches 1682; Conservative 0; Mismatches 995; Indels 69; Gaps 6;

 Qy 12 GCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGCTTGGCTCCGGCTCGGG 71
 || || || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGGCCAGCGGACTGGCTCCGCCGCCAAGA 125

 Qy 72 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 131
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 126 TGATGACTTTTCATGAACCTCCAGAACTTTCCCTCTGATCCACCTGAGCCTCTGCC 185

 Qy 132 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCAGTGCTGCTTGT 191
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 186 ACATTTCTTATTGAGCCTGAAGAAGCTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

 Qy 192 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCAAGTGCAACGGGAGTGGGTGCCCA 251
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 246 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTCAAGTGTAAAGTGAATGGGTTCATCA 305

 Qy 252 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCACCATGGAGGT 311
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 306 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGCTCATGTCCGGGAAGT 365

 Qy 312 CCGCATTAAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAGGAATACTGGTG 371
 . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 366 GAGCATTGAGATTCGCGCCAGCAAGTGGAAAGAACTTTGGACCTGAAGATTACTGGTG 425

 Qy 372 CCAGTGCCTGGCATGGAGCTCCTCGGGACCAAGAGTCAGAAGGCCTACATCCGCAT 431
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 426 CCAGTGTGTGGCCTGGAGCTCCGGTACACAAAGAGGCCAGGGCTATGTGCGCAT 485

Qy	432	AGC-----CAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGCCAAGGAGGTGTCCT	485
Db	486	TGCATNNNNNNNNNNNCGGAAGACATTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTT	545
Qy	486	GGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA	545
Db	546	GGAACAGGAAGTCTTACTCCAGTGTGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGA	605
Qy	546	GTGGCTCCCGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACTCACCGC	605
Db	606	ATGGTTGAAAAATGAAGACATAATTGATCCCGTTGAAGAGATCGGAATTTTATATTACTAT	665
Qy	606	GGAGCACAGCCTGGTGGTGCACAGGCCCTTGCTGACACGCCAATCACACCTGCGT	665
Db	666	TGATCACAAACCTCATCATAAAGCAGGCCGACTCTGTGATACTGCAAATTACACCTGTGT	725
Qy	666	GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTATCGTCTACGTGAA	725
Db	726	TGCCAAAACATTGTTGCCAAGAGGAAAAGTACAACGTCCACTGTGATAGTCTATGTCAA	785
Qy	726	CGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGCGCTG	785
Db	786	CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGGGTA	845
Qy	786	GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTCTGTGA	845
Db	846	TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGTGCCTTCTGTGA	905
Qy	846	GGGGCAGAATGTCCAGAAAACAGCCTGCCACCCCTGTGCCAGTAGACGGCAGCTGGAG	905
Db	906	AGGGCAGAGTGTGAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGAC	965
Qy	906	CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGAGCCGTGAGTG	965
Db	966	GCCATGGAGCAAGTGGTCACTTGTGGAAGTGCAGCTGGCAGGGACTGACCTGGACACCCG	1025
Qy	966	CTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTGGACACCCG	1025
Db	1026	CACGGCGCCAGCCCCAAGAATGGAGGCAAGGACTGGACGGCCTCGTCTGCAATCAA	1085
Qy	1026	CAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGGCCCTCTA	1085
Db	1086	GAAC TGCACTGATGGCTTGCATGCAGACTGCTCTGATTGAGATGATGTTGCTCTCTA	1145
Qy	1086	TGTGGG---CCTCATGCCGTGCCGTGCCTGGCCTGCTGCTGCTGCTGCTCATCCT	1142
Db	1146	TGTTGGGATTGTGATAGCAGTGATCGTTGCCCTGGCGATCTCTGAGTTGCTGCTTGT	1205
Qy	1143	CGTTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC	1202
Db	1206	TGTGTATCGGAAGAACATCGTGACTTGAGTCAGATATTATTGACTCTCGGCACCAA	1265
Qy	1203	CTCAGGCTCCAGCCGTAGCATCAAGCCAGCAAAGCAGACAACCCCATCTGCTCAC	1262
Db	1266	TGGGGCTTCAGCCTGTGAACATCAAG-----GCAGCAAGACAAGAGATCTGCTGGC	1316

Db	2157	TGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTTCAATTAAAGGCAG	2216
Qy	2112	TTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT	2171
Db	2217	CACCCACAAACCTGCGCCTGTCAATTACGATATCGCCATTCCCTGTGGAAGAGCAAATT	2276
Qy	2172	CCTTGTCACTACCAGGAGATCCCCTTTATCACATCTGGAATGGCACGCAGCGGTACTT	2231
Db	2277	GCTGGCTAAATATCAGGAAATTCCATTTCACCATGTTGGAGTGGATCTCAAAGAAACCT	2336
Qy	2232	GCACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT	2291
Db	2337	GCACTGCACCTTCACCTCTGGAAAGATTAGCCTGAACACAGTGGAGCTGGTTGCAAAC	2396
Qy	2292	GTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTAACATCACCAA	2351
Db	2397	CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCAGCTCAACTGCACCGTGTCA	2456
Qy	2352	GGACACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTCCCAGCCCTGGT	2411
Db	2457	GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTCGAACACCATACCACGGTCAC	2516
Qy	2412	GGGCCAGTGCCTCAAGATCCCCTTCATTGGCAGAAGATAATTCCAGCCTGGA	2471
Db	2517	GGGGCCCAGTGCTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA	2576
Qy	2472	CCCACCCCTGTAGGCAGGGTGCCACTGGGGACTCTGGCCAGAAACTCCACCTGGACAG	2531
Db	2577	TGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCATAAGCTGAACCTGGACAG	2636
Qy	2532	CCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTAACCTGTGGA	2591
Db	2637	GTACTTGAATTACTTGCCACCAAATCCAGCCAACTGGCGTAATCCTGGATCTTGGA	2696
Qy	2592	GGCGCGGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGG	2651
Db	2697	AGCACAGAACTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTGGAAAGAAATGGG	2756
Qy	2652	CCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGTGA	2697
Db	2757	AAGACATGAAACGGTGGTGTCCCTAGCAGCAGAAGGGCAGTATTAA	2802

RESULT 2

AY406493

LOCUS AY406493 2791 bp DNA linear GSS 15-DEC-2003
 DEFINITION Mus musculus UNCSC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY406493
 VERSION AY406493.1 GI:39762467
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2791)

Db	518	AGGAACCCTGGGAAAGGAAGTGTCCCTGGAGCAGGAAGTCTTACTCCAGTGTGGCCAC	577
Qy	518	CGGAGGGCATCCCTCCAGCCAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT	577
Db	578	CTGAAGGGATCC-----GTAGAATGGCTAAAGAATGAAGACATAATTGATCCTG	626
Qy	578	CCCTGGACCCCAATGTATAACATCACGCCGGAGCACAGCCTGGTGGTGCACAGGCCGCC	637
Db	627	CTGAAGATCGAACCTTTATATTACTATCGATCACAAACCTGATCATCAAGCAAGCCCCGAC	686
Qy	638	TTGCTGACACGGCAAACACACCTGCGTGGCAAGAACATCGTGGCACGTCGCCGCCAGCG	697
Db	687	TCTCAGATAACAGCAAATTATAACCTGTGTTGCCAAAATATTGTTGCCAAGAGAAAAAGCA	746
Qy	698	CCTCCGCTGCTGTACCGTACCGTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCG	757
Db	747	CCACAGCCACTGTACCGTGTATGTTATGGTGGCTGGCACCTGGACAGAGTGGTCTG	806
Qy	758	TCTGCAGCGCCAGCTGTGGCGCGCTGGCAGAAACGGAGGCCGGAGCTGCACCAACCCGG	817
Db	807	TGTGTAACAGCCGCTGTGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAG	866
Qy	818	CGCCTCTAACGGGGCGCTTCTGTGAGGGCAGAACATGTCCAGAAAACAGCCTGCGCCA	877
Db	867	CCCCACTCAATGGTGGGCGCTTCTGTGAGGGCAGAGTGTGCAGAAAATAGCATGCACTA	926
Qy	878	CCCTGTGCCAGTAGACGGCAGCTGGAGCCCCTGGAGCAAGTGGTCGGCTGTGGCTGG	937
Db	927	CGTTATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTG	986
Qy	938	ACTGCACCCACTGGCGAGCCGTGAGTGTCTGACCCAGCACCCCCGCAACGGAGGGGAGG	997
Db	987	AATGCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAAGCCCCCAAGAACGGGGTAAGG	1046
Qy	998	AGTGCAGGGCACTGACCTGGACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTG	1057
Db	1047	ACTGTGATGGCCTGGTCCTCAATCCAAGAACTGCACTGATGGCTGTGCATGCAGGCTG	1106
Qy	1058	CTTCTGCCCTGAGGACGTGCCCTCTATGTGGG---CCTCATGCCGTGGCGTCTGCC	1114
Db	1107	CTCCTGACTCAGATGATGTGGCTCTACGTGGGATTGTGATCGCTGTAACAGTCTGTC	1166
Qy	1115	TGGCCTGCTGCTGCTTCCTCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACT	1174
Db	1167	TGGCGATCACTGTTGTTGCCCTGTTGTGATCGGAAGAACCAACCGTGACTTTGAGT	1226
Qy	1175	CAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTCCAGCCCCGTACGCATCAAGCCCA	1234
Db	1227	CTGACATCATTGACTCCTCAGCACTCAATGGCGGTTTCAGCCTGTGAACATCAAG---	1282
Qy	1235	GCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCAACCAACCA	1294
Db	1283	-----GCTGCCAGACAAGATCTCCTGGCTGTCCCCCTGACCTCACCTCAGCTGCAGCCA	1337
Qy	1295	CCTACCAGGGCAGTCTGTCCCCGGCAGGATG-----GGCCAGCCCCAAGTTCCAGCT	1349

Db	1338	TGTACAGGGGACCTGTCTATGCTCTGCATGATGTCAGACAAAATCCCAATGACCAACT	1397
Qy	1350	CACCAATGGGCACCTGCTCAGCCCCCTGGTGGCGGCCACACACTGCACCACAGCTC	1409
Db	1398	CTCCAATTCTGGACCCACTACCCAACTTGAAAATCAAAGTGACAACAGCTCAGGTGCTG	1457
Qy	1410	TCCCACCTC---TGAGGCCGAGGAGTCGTCTCCGCCTCTCCACCC-----	1453
Db	1458	TCACTCCTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCAGATGACCCAGT	1517
Qy	1454	-----AGAACTACTTCCGCTCCCTGCCCGAGGCA	1483
Db	1518	CCTTGCTAGAGAATGAGGCCCTAACCTGAAGAACCAAGAGCCTCGCAAGACAGACTGACC	1577
Qy	1484	CCAGCAACATGACCTATGGGACCTTCAACTTCCCTGGGGGCCGGCTGATGATCCCTAATA	1543
Db	1578	CATCCTGCACAGCATTGGTACCTCAACTCTCTGGGGTACCTCATCATTCTTAATT	1637
Qy	1544	CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCT	1603
Db	1638	CAGGAGTAAGCTTGCTGATTCCCGCTGGGCCATTCCCTCAGGGAGAGTCTATGAAATGT	1697
Qy	1604	ACCTCACGCTGCACAAGCCGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGC	1663
Db	1698	ATGTGACTGTACACAGGAAAGAAAATATGAGGCCCATGGAAGACTCTCAGACCCCTAC	1757
Qy	1664	TGAGTCCCATCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGG	1723
Db	1758	TTACCCCTGTGGTGAGCTGTGGCCTCTGGAGCTCTGCTGACCCGCCGTACCTCA	1817
Qy	1724	CTATGGACCCTGTGGGAGGCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAGCAGT	1783
Db	1818	CTCTGCATCACTGTGCAGACCCAGCACCGAGGACTGGAAGATCCAGCTCAAAACCAGG	1877
Qy	1784	CGTGCAGGGCAGCTGGAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCT	1843
Db	1878	CAGTGCAGGGACAATGGGAGGATGTTGTGGTGGTGGGAGGAGAACTCACAACCCCT	1937
Qy	1844	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTTG	1903
Db	1938	GTTACATTCACTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATG	1997
Qy	1904	CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTG	1963
Db	1998	CCCTGGTTGGCAGTCCACCAACAAAGCAGCTGCAAGCGCTTAAACTGGCCATCTTG	2057
Qy	1964	CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGCTACTGCCTGCATGACACCC	2023
Db	2058	GGCCCCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCACTGCCTGGATGACACAC	2117
Qy	2024	ACGATGCCACTCAAGGAGGTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGG	2083
Db	2118	AGGATGCCCTGAAGGAAGTTCTACAACCTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAG	2177
Qy	2084	AGCCACGGGTCTGCACCTCAAGGACAGTTACCACAAACCTGCGCCTATCCATCCACGATG	2143
Db	2178	AACCCAAGGCTCTCATTAAAGGCAGCATTCCACAAACCTGCGCCTGTCTATTGACA	2237

RESULT 3

AK031655

LOCUS AK031655 3790 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030473H24 product:unc5 homolog (C. elegans) 3, full insert sequence.
ACCESSION AK031655
VERSION AK031655.1 GI:26327502
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3790)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.
URL:<http://genome.gsc.riken.jp/>
URL:<http://fantom.gsc.riken.jp/>.

FEATURES	Location/Qualifiers
source	1. .3790 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:6030473H24" /db_xref="taxon:10090" /clone="6030473H24" /sex="male" /tissue_type="testis" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="13 days embryo"
CDS	118. .2970 /note="unnamed protein product; putative unc5 homolog (C. elegans) 3 (MGD MGI:1095412, GB NM_009472, evidence: BLASTN, 99%, match=464)" /codon_start=1 /protein_id="BAC27495.1" /db_xref="GI:26327503" /translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ KDHVVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTSKRKAYV RIAYLRKTFEQEPLGKEVSLEQEVLQLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRSTTATVIVYVNGGWSTWTEWSVCNSRC GRGYQKRRTCTNPAPLNNGAFCEGQSQKIACTLCPVDGRWTWSKWSTCGTECH WRRRECTAPAKPNNGKDCDGVLVLSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA RQDLLAVPPDLTSAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT PQDDLAEFSSKLSPQMTQSLLNEALNLKNQSLARQTDPSCTAFGTFSNLGGHLIIPN SGVSLLIPAGAIQPGGRVYEMYVTVHRKENMRPPMEDSQTLITPVVSCGPPGALLTRPV ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHILTEN LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM GGQLLEEPKALRFKGSIHNLRLSIHDIASLWKSLLAKYQEIPFYHIWSGSQRNLHC TFTLERLSSLNTVELVKLCVQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVT GPSAFSIPLPIRQKLCSSLDAPOTRGHDRWMLAHKLNLDRLNYFATKSSPTGVILDWEAQNFPDGNLSMLAAVLEEMGRHETVVLAAEGQY"

ORIGIN

```

Query Match           34.2%; Score 923.4; DB 3; Length 3790;
Best Local Similarity 61.3%; Pred. No. 1.4e-186;
Matches 1659; Conservative 0; Mismatches 926; Indels 120; Gaps 6;

    104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGGCCGAGGATGTGTACA 163
          ||||| ||| | ||||| | | ||| ||||||| | ||| ||| ||| | | ||| |
    275 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGGCCGAGGAAGCTTACA 334

    164 TCGTCAAGAACAAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCT 223
          ||| ||||||| | ||| | | ||| | ||| | ||| ||| ||| ||| |
    335 TTGTGAAGAACAAAGCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 394

```

Qy	224	TCAAGTGCACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG	283
Db	395	TCAAGTGCACAGCGAGTGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG	454
Qy	284	GGAGCAGTGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGA	343
Db	455	AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTACGCCAGCAGGTGGAGG	514
Qy	344	AGGTGTTGGGCTGGAGGAATACTGGTGCAGTGCAGTGGCATGGAGCTCCTCGGGCACCA	403
Db	515	AACTGTTGGGCTGAAGATTACTGGTGCAGTGTGTGGCCTGGAGCTCAGCAGGACTA	574
Qy	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGC	463
Db	575	CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTGAGCAGGAAC	634
Qy	464	CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGCCACCGGAGG	523
Db	635	CCTTGGGAAAGGAAGTGTCCCTGGAGCAGGAAGTCTTACTCCAGTGTGGCCACCTGAAG	694
Qy	524	GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG	583
Db	695	GGATCCCAGTGGCTGAGGTAGAATGGCTAAAGAACATGAAGACATAATTGATCCTGCTGAAG	754
Qy	584	ACCCCAATGTATACATCACGGGAGCACAGCCTGGTGGCGACAGGCCGCCTGCTG	643
Db	755	ATCGGAACCTTATATTACTATCGATCACAACTGATCATCAAGCAAGGCCGACTCTCAG	814
Qy	644	ACACGGCCAACCTACACCTGCGTGGCAAGAACATCGTGGCACGTCGCCAGCGCCCTCG	703
Db	815	ATACAGCAAATTATACCTGTGTTGCCAAAATATTGTTGCCAAGAGAAAAAGCACACAG	874
Qy	704	CTGCTGTCATCGTCTACGTGAACGGTGGTGGTCACGTGGACCGAGTGGCTCGTCA	763
Db	875	CCACTGTCATCGTGTATGTTAATGGTGGCTGGCACCTGGACAGAGTGGCTGTGTGTA	934
Qy	764	GCGCCAGCTGTGGCGCGCTGGCAGAACGGAGCCGGAGCTGCACCAACCGGCCCTC	823
Db	935	ACAGCCGCTGTGGCGAGGATATCAGAACGACAAGAACCTGCACCAACCCAGCCCCAC	994
Qy	824	TCAACGGGGCGCTTCTGTGAGGGGCAGAACATGCCAGAACAGCCTGCCACCCCTGT	883
Db	995	TCAATGGTGGGCCCTCTGTGAGGGCAGAGTGTGCAGAAAATAGCATGCACACGTTAT	1054
Qy	884	GCCCAGTAGACGGCAGCTGGAGCCGTGGAGCAAGTGGTGGCCTGTGGCTGGACTGCA	943
Db	1055	GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGACTGAATGCA	1114
Qy	944	CCCACGGCGAGCCGTGAGTGCCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCC	1003
Db	1115	CCCACGGCGAGGAGGGAGTGTACAGCACCAAGAACGGGGTAAGGACTGTG	1174
Qy	1004	AGGGCACTGACCTGGACACCCGCAACTGTACCGTAGCCTGTGTACACAG-----	1055
Db	1175	ATGGCCTGGTCTCCAATCCAAGAACATGCACGTGACTGATGGCTGTGCATGCAGGGATTCA	1234

Qy	1056	-----TGCTTCTGGCC	1066
Db	1235	ACCCCATTCAACTGAGCACAGACCCCAGAATGAATATGGATTTCTCTGCTCCTGACT	1294
Qy	1067	CTGAGGACGTGGCCCTCTATGTGGG---CCTCATGCCGTGGCCGTCTGCCTGGCCTGC	1123
Db	1295	CAGATGATGTGGCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGATCA	1354
Qy	1124	TGCTGCTTGTCCCTCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGG	1183
Db	1355	CTGTTGTGGTGGCCCTGTTGTATCGGAAGAACCAACCGTGACTTGAGTCTGACATCA	1414
Qy	1184	CTGACTCGTCCATTCTCACCTCAGGCTTCAGGCCGTCAAGCCCAGCAAAGCAG	1243
Db	1415	TTGACTCCTCAGCACTCAATGGCGGCTTCAGCCTGTGAACATCAAG-----GCTG	1465
Qy	1244	ACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCAACCTACCAAGG	1303
Db	1466	CCAGACAAAGATCTCCTGGCTGTCCCCCTGACCTCACCTCAGCTGCAGCCATGTACAGGG	1525
Qy	1304	GCAGTCTCTGTCCCCGGCAGGATG----GGCCCAGCCCCAAGTTCCAGCTCACCAATGG	1358
Db	1526	GACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCAATGACCAACTCTCCAATTC	1585
Qy	1359	GCACCTGCTCAGCCCCCTGGGTGGCGGCCACACACTGCACCAACAGCTCTCCCACCTC	1418
Db	1586	TGGACCCACTACCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCC	1645
Qy	1419	----TGAGGCCGAGGAGTTCGTCTCCGCCTCTCCACCC-----	1453
Db	1646	AGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCAGATGACCCAGTCCTTGCTAG	1705
Qy	1454	-----AGAACTACTTCCGCTCCCTGCCCGAGGCACAGCAACA	1492
Db	1706	AGAATGAGGCCCTAACCTGAAGAACAGAGCCTCGAAGACAGACTGACCCATCTGCA	1765
Qy	1493	TGACCTATGGGACCTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGGTATCA	1552
Db	1766	CAGCATTGGTACCTCAACTCTTGGGGTCACCTCATCTTCTAACCTAATTCAAGGAGTAA	1825
Qy	1553	GCCTCCTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGC	1612
Db	1826	GCTTGCTGATTCCCGCTGGGCCATTCTCAGGGGAGAGTCTATGAAATGTATGTGACTG	1885
Qy	1613	TGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCA	1672
Db	1886	TACACAGGAAAGAAAATATGAGGCCCTCATGGAAGACTCTCAGACCCCTACTACCC	1945
Qy	1673	TCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACC	1732
Db	1946	TGGTGAGCTGTGGCCTCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGC	2005
Qy	1733	ACTGTGGGGAGCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAGCAGTCGTGCGAGG	1792
Db	2006	ACTGTGCAGACCCCAACGACCCGAGGACTGGAAGATCCAGCTCAAAACCAGGCAGTGC	2065
Qy	1793	GCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTACTACTGCC	1852

Db	2066	GACAATGGGAGGATGTTGTGGTGGTGGGAGGAACCTCACAACCCCCTGTTACATT C	2125
Qy	1853	AGCTGGAGGCCAGTGCTGCTACGTCTCACCGAGCAGCTGGGCCGCTTGCCCTGGTGG	1912
Db	2126	AGCTGGATGCAGAGGCTTGCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTG	2185
Qy	1913	GAGAGGCCCTCAGCGTGGCTGCCGCCAAGGCCTCAAGCTGCTCTGTTGCGCCGGTGG	1972
Db	2186	GGCAGTCCACCACCAAAGCAGCTGCCAACGCTTAAACTGCCATCTTGGCCCTCT	2245
Qy	1973	CCTGCACCTCCCTCGAGTACAACATCCGGTCACTGCCTGCATGACACCCACGATGCAC	2032
Db	2246	GCTGCTCTCCCTGGAGTACAGCATTAGAGTCACTGCCTGGATGACACACAGGATGCC	2305
Qy	2033	TCAAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGGCCACGGG	2092
Db	2306	TGAAGGAAGTTCTACAACCTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGG	2365
Qy	2093	TCCTGCACTTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCT	2152
Db	2366	CTCTCGTTTAAAGGCAGCATCCACAAACCTGCGCTGTCTATTGACATGCCATT	2425
Qy	2153	CCCTGTGGAAGAGTAAGCTCCTGTCAGCTACCAGGAGATCCCCTTTATCACATCTGGA	2212
Db	2426	CCCTCTGGAAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTACACATCTGGA	2485
Qy	2213	ATGGCACGCAGCGGTACTTGCAC TGACCTTCACCCCTGGAGCGTGTAGCCCCAGCACTA	2272
Db	2486	GTGGCTCTCAAAGAACCTCCACTGCACCTCACTCTGGAAAGACTCAGCCTAACACAG	2545
Qy	2273	GTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCA	2332
Db	2546	TGGAACTGGTTGCAAACCTGTGTGCGGCAGGTTGAAGGAGAACGGCAGATCTCCAGC	2605
Qy	2333	TCAACTTCAACATACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAG	2392
Db	2606	TCAACTGTACTGTGTCAAGAGAACCTACTGGCATCGACTTACCTCTCCTGGACCCCTGCTA	2665
Qy	2393	CGGGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGA	2452
Db	2666	GTACCATCACCACGTGCAACGGACCAAGTGCTTTCAGCATTCCCTCCCTATCCGGCAGA	2725
Qy	2453	AGATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCTGGCGACTGGCGACTCTGGCC	2512
Db	2726	AGCTATGCAGCAGCCTGGATGCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCC	2785
Qy	2513	AGAAAATCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAGCCA	2572
Db	2786	ATAAAATCAACCTGGACAGGTACTGAATTACTTGCCACCAATCGAGCCAACTGGCG	2845
Qy	2573	TGATCCTCAACCTGTGGAGGCCGGCAGCTTCCCCAACGGCAACCTCAGCCAGCTGGCTG	2632
Db	2846	TAATCCTGGATCTTGGAGACAGAACCTCCCAGATGGAAACCTGAGCAGCTGGCAG	2905
Qy	2633	CAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGT	2692

Db 2906 CCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTACTTGGCAGCAGAAGGACAGT 2965
Qy 2693 GCTGA 2697
 |||
Db 2966 ATTGA 2970

RESULT 4

CR598115

LOCUS CR598115 1852 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CS0DA006YG16 of Neuroblastoma of Homo sapiens (human).
 ACCESSION CR598115
 VERSION CR598115.1 GI:50478922
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1852)
 AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
 <http://fulllength.invitrogen.com/> InVitroGen Corporation 1600
 Faraday Avenue
 REFERENCE 2 (bases 1 to 1852)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES Location/Qualifiers
 source 1. .1852
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA006YG16"
 /tissue_type="Neuroblastoma"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 32.4%; Score 874; DB 3; Length 1852;
Best Local Similarity 100.0%; Pred. No. 4.8e-176;
Matches 874; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1884 CGAGCAGCTGGGCCGCTTGCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 1943

Db	61	CGAGCAGCTGGCCGCTTGCCTGGTGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG	120
Qy	1944	CCTCAAGCTGTTCTGTTGCCTGGCTGCACCTCCCTCGAGTACAACATCCGGT	2003
Db	121	CCTCAAGCTGTTCTGTTGCCTGGCTGCACCTCCCTCGAGTACAACATCCGGT	180
Qy	2004	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAACAGCT	2063
Db	181	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAACAGCT	240
Qy	2064	GGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCACCTCAAGGACAGTTACCACAAACCT	2123
Db	241	GGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCACCTCAAGGACAGTTACCACAAACCT	300
Qy	2124	GCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTGTCAGCTA	2183
Db	301	GCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCCTGTCAGCTA	360
Qy	2184	CCAGGAGATCCCCTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACCTGCACCTT	2243
Db	361	CCAGGAGATCCCCTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACCTGCACCTT	420
Qy	2244	CACCCGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	2303
Db	421	CACCCGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	480
Qy	2304	GGTGGAGGGCGACGGCAGAGCTTCAGCATCAACATCACCAAGGACACAAGGTT	2363
Db	481	GGTGGAGGGCGACGGCAGAGCTTCAGCATCAACATCACCAAGGACACAAGGTT	540
Qy	2364	TGCTGAGCTGGCTCTGGAGAGTGAAAGCGGGGTCCCAGCCCTGGTGGCCCCAGTGC	2423
Db	541	TGCTGAGCTGGCTCTGGAGAGTGAAAGCGGGGTCCCAGCCCTGGTGGCCCCAGTGC	600
Qy	2424	CTTCAAGATCCCCTTCCTCATTGGCAGAAGATAATTCCAGCCTGGACCCACCTGTAG	2483
Db	601	CTTCAAGATCCCCTTCCTCATTGGCAGAAGATAATTCCAGCCTGGACCCACCTGTAG	660
Qy	2484	GCGGGGTGCCGACTGGCGACTCTGGCCAGAAACTCCACCTGGACAGCCATCTCAGCTT	2543
Db	661	GCGGGGTGCCGACTGGCGACTCTGGCCAGAAACTCCACCTGGACAGCCATCTCAGCTT	720
Qy	2544	CTTTCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGAGGCCGGCACTT	2603
Db	721	CTTTCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGAGGCCGGCACTT	780
Qy	2604	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGCCAGCCAGACGC	2663
Db	781	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGCCAGCCAGACGC	840
Qy	2664	TGGCCTCTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	841	TGGCCTCTCACAGTGTGGAGGCTGAGTGCTGA	874

RESULT 5
AK018177

LOCUS AK018177 3866 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male medulla oblongata cDNA, RIKEN full-length enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION AK018177
VERSION AK018177.1 GI:12857775
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3866)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,

Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGGATCCAAGAGCTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAAATTAAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES Location/Qualifiers

source 1. .3866
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6330415E02"
/db_xref="taxon:10090"
/clone="6330415E02"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS 417. .3254
/note="unnamed protein product; TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus] (SPTR|O08722, evidence: FASTY, 96.5%ID, 100%length, match=2835)
putative"
/codon_start=1
/protein_id="BAB31108.1"
/db_xref="GI:12857776"
/translation="MRARSGVRSALLLALLCWDPTPSLAGVDSAGQVLPDPSYPSAPA
EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG
LRVREVQIEVSQQVEELFGLEDWCQCVAWSSSGTTSRRAYIRIAYLRKNFQEQPL
AKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS
DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAESPCSNRCGRGWQKRRTCTNP
APLNGGAFCEGQAFQKTACTTVCVPVDGAWEWSKWSACSTECAHWRSCRECMAPPQNG
GRDCSGTLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVAVFVVVAVLM

AEGVIVYRRNCRDFDTIDTSSAALTGGFPVNFKTARPNNQLLHPSAPPDLTASAG
IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSТИGSGSLADGADLLGVLP
GTYPGDFSRDTHFLHLRSASLGQSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTVSLL
VPNGAIHQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCPTGLLLCRPVVLTVPH
CAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF
MGEYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE
EPKPLLFDKSYHNLRLSLHDIPIAHWRSKLLAKYQEIPFYHWNGSQRALHCTFTLER
HSLASTEFTCKVCVRQVEGEGQIFQLHTTAAETPAGSLDALCSAPGNAITTQLGPYAF
KIPLSIRQKICSSLADPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ
QDDGDLNSLASALEEMGKSEMLVAMATDGDC"

ORIGIN

Query Match 32.3%; Score 871.4; DB 3; Length 3866;
 Best Local Similarity 60.6%; Pred. No. 1.9e-175;
 Matches 1645; Conservative 0; Mismatches 941; Indels 129; Gaps 8;

Qy	112	GCCAACCCGGACCTGCTCCCCACTTCCCTGGGGAGGCCGAGGATGTGTACATCGTCAAG	171
Db	540	GCCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG	599
Qy	172	AACAAGCCAGTGCTGTTGTGCAAGGCCGTGCCGCCACGCAGATCTCTTCAGTGC	231
Db	600	AACAAGCCAGTGGAACTGCACTGCAGAGCCTCCCTGCCACGCAGATCTACTTCAGTGT	659
Qy	232	AACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT	291
Db	660	AATGGCGAGTGGGTGAGCCAGAATGACCACGTACACAGGAGAGCCTGGATGAGGCCACA	719
Qy	292	GGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTT	351
Db	720	GGCTTGGGGTGCGAGAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACCTTC	779
Qy	352	GGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCCGGGCACCACCAAGAGT	411
Db	780	GGGCTCGAGGACTACTGGTGCCAGTGCCTGGCCTGGAGCTCTCGGAACTACCAAGAGT	839
Qy	412	CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGGCCCTGGCC	471
Db	840	CGCCGAGCCTACATCCGCATTGCCCTACCTGCGCAAGAACCTTGACCAGGAGCCTCTGGCC	899
Qy	472	AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCT	531
Db	900	AAGGAGGTACCCCTGGATCATGAGGTCCCTCTGCAGTGCCGCCACCGGAGGGAGTGCCT	959
Qy	532	CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT	591
Db	960	GTGGCTGAGGTGGAATGGCTCAAGAACATGAAGATGTCATTGACCCGCTCAGGACACTAAC	1011
Qy	592	GTATACATCACGGGGAGCACAGCCTGGTGGCGACAGGCCGCCCTGCTGACACGGCC	651
Db	1020	TTCCTGCTACCATTGACCACAAACCTCATCATCCGCCAGGCGCCTCTCAGACACGGCC	1071
Qy	652	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTGCCGCCAGGCCCTCCGCTGCTGTC	711
Db	1080	AACTACACCTGTGTGGCCAAGAACATCGTGGCCAAGCGCCGGAGCACCGCGGCCACAGTC	1133
Qy	712	ATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGCCGTCTGAGCGCCAGC	771

Db	1140	ATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCGTTCAAATCGC	1199
Qy	772	TGTGGCGCGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGG	831
Db	1200	TGTGGCGAGGCTGGCAGAACGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGA	1259
Qy	832	GGCGCTTTCTGTGAGGGCAGAATGTCAGAAAAACAGCCTGCCACCCGTGCCAGTA	891
Db	1260	GGCGCCTTCTGTGAGGGACAGGCCCTCAGAAGACAGCTTGACCAACCACCGTGTGCCAGTG	1319
Qy	892	GACGGCAGCTGGAGCCC GTGGAGCAAGTGGTCGGCTGTGGCTGGACTGCACCCACTGG	951
Db	1320	GATGGAGCGTGGACCGAGTGGAGCAAGTGGCTGCCCTGCAGCACAGAGTGTGCGCACTGG	1379
Qy	952	CGGAGCCGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGGCACT	1011
Db	1380	CGCAGCCCGAGTGCATGGCACCGCCACCCAGAACGGAGGCCGTGACTGCAGCGGGACG	1439
Qy	1012	GACCTGGACACCCGCAACTGTACCAAGTGCACCTCTGTGTACACAGTGCTTCTGGCC-----	1066
Db	1440	CTACTTGACTCCAAGAACTGCACTGATGGCTGTGCGTGCTGAATCAGAGAACTCTAAC	1499
Qy	1067	-----CTGAGGACGTGGCCCTCTATGTGGCCCTCATC	1098
Db	1500	GACCTAAAAGCCACCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCCTGTG	1559
Qy	1099	GCCGTGGCCGTCTGCCTGGCTCTGCTGCTGCTGTC---CTCATCCTCGTTATTGCCGG	1155
Db	1560	GTGGCCGTCTTGTTGGTAGCGGTTCTCATGGCCGAGGGAGTGATCGTATACCGGAGA	1619
Qy	1156	AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGT---CCATTCTCACCTCAGGCTTC	1212
Db	1620	AACTGCCGGACTTCGACACGGACATCACCAGCTCTGCCGCCCCCTCACTGGTGCTTC	1679
Qy	1213	CAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCATCTGCT-----CACCATC	1266
Db	1680	CACCTGTCAACTCAAGACTGCAAGGCCAACAACCCGCAGCTCCTGCACCCGTCCGCC	1739
Qy	1267	CAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCCCGGCAGGA-	1325
Db	1740	CCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCTGTGTATGCCCTGCAGGAC	1799
Qy	1326	-----TGGGCCAGCCCCAAG-----	1341
Db	1800	TCCGCCGACAAGATCCCCATGACTAATTGCCCTGCTGGATCCCCCTGCCAGCCTCAAG	1859
Qy	1342	TTCCAGCTACCAATGGCACCTGCTCAG-----	1371
Db	1860	ATCAAGGTCTATAACTCCAGCACCATCGTTCTGGGTCTGCCCTGGCTGATGGAGCCGAC	1919
Qy	1372	CCCCTGGGTGGCGGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCGAGGAG	1431
Db	1920	CTGCTGGGTGTCTCCCGCCGGCACGTACCCAGGCATTTCTCCGGACACCCATTTC	1979
Qy	1432	TTCGTCTCCCGCCTCTCCACCC-----AGAACTACTCCGCTCCCTGCCCGAGGCACC	1485

Db	1980	CTGCACCTGCGCAGTGCCAGCCTGGTCCAGCACCTCCTGGCCTACCTCGGGACCCC	2039
Qy	1486	AGCAACATGACCTATGGACCTCAACTTCTCGGGGCCGGCTGATGATCCCTAACATA	1545
Db	2040	AGCAGCAGTGTCAAGCGCACCTTGTTGCCCTGGGAGGAAGGCTGAGCCTCCCCGGCACA	2099
Qy	1546	GGTATCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTAC	1605
Db	2100	GGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTAT	2159
Qy	1606	CTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCTGCTG	1665
Db	2160	CTACATATCAACAAGGCCAAGCACCCCTCCACTTCAGAAGTTCCCAGACAGTATTG	2219
Qy	1666	AGTCCCATCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT	1725
Db	2220	AGCCCCCTCGGTGACCTGTGGGCCACAGGGCTACTCCTGTGCCGCCCTGTCGTCCCTCACC	2279
Qy	1726	ATGGACCACTGTGGGAGGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCG	1785
Db	2280	GTGCCCTACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTCAGCTCAAGACCCAGGCC	2339
Qy	1786	TGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGGCCCTCCCACCTCTAC	1845
Db	2340	CATCAGGGCCACTGGGAGGAGGTGGTGACCTGGATGAGGAGACCCCTAACACACCCCTGC	2399
Qy	1846	TACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCC	1905
Db	2400	TACTGCCAGCTGGAGGCTAAGTCTGCCACATCCTGCTGGACCAGCTGGGTCCTACGTA	2459
Qy	1906	CTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAACGGCCTCAAGCTGCTCTGTTGCG	1965
Db	2460	TTCATGGCGAGTCTACTCTCGCTCTGCAGTCAAGCGGCTCAGCTGGCCATCTCGCC	2519
Qy	1966	CCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCAC	2025
Db	2520	CCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGCTACTGTCTGGAGGACACACT	2579
Qy	2026	GATGCACTCAAGGAGGTGGTGAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAG	2085
Db	2580	GTAGCACTGAAGGAGGTCTGGAGCTGGAGAGGACTCTGGTGCTACTTGGTGGAGGAG	2639
Qy	2086	CCACGGGTCTGCACCTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTG	2145
Db	2640	CCCAAGCCTTGTCTTTAAGGACAGTTACCACAAACCTACGCCCTCCCTCCATGACATC	2699
Qy	2146	CCCAGCTCCCTGTGGAAGAGTAAGCTCCTGTCAGTACCAAGGAGATCCCCTTTATCAC	2205
Db	2700	CCCCATGCCCACTGGAGGAGCAAACACTGGCCAAGTACCAAGGAGATTCCCTCTACAC	2759
Qy	2206	ATCTGGAATGGCACGCAGCGTACTTGCAGTCAGTACCAAGGAGATCCCCTTTATCAC	2265
Db	2760	GTCTGGAATGGCAGCCAGAGAGGCCCTGCAGTCAGTACCTTCACCCGGAGAGGCATAGCCTG	2819
Qy	2266	AGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGC	2325
Db	2820	GCCTCCACGGAGTTCACCTGTAAGGTCTGCGTGCAGGTCGAAGGGGAAGGCCAGATT	2879

RESULT 6

AY406492

LOCUS AY406492 2802 bp DNA linear GSS 15-DEC-2003
 DEFINITION Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY406492
 VERSION AY406492.1 GI:39762466
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 2802)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 2802)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1. .2802
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"

gene <1. .>2802
 /gene="UNC5C"
 /locus_tag="HCM2575"

ORIGIN

Query Match	30.2%	Score 814;	DB 9;	Length 2802;
Best Local Similarity	53.8%	Pred. No. 3.4e-163;		
Matches 1477;	Conservative 0;	Mismatches 1200;	Indels 69;	Gaps 6;

```

Qy      12 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGGCTCGGG 71
        ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCAAGA 125
Qy      72 TGCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTCCAACCCGGACCTGCTTCC 131
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      126 TGATGACTTTTTCATGAACCTCCCAGAAACTTTCCCTCGGATCCACCTGAGCCTTGCC 185
Qy      132 CCACTCCTGGTGGAGGCCGAGGATGTGTACATCGTAAGAACACAAGCCAGTGTGCTTGT 191
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      186 ACATTTCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245
Qy      192 GTGCAAGGCC-----GTGCCCGCACGCAGATCTTCTCAAGTGCAACGGGAGTGGGT 245
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      246 CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 305
Qy      246 GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCCACCAT 305
Db      306 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 365
Qy      306 GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAGGAATA 365
Db      366 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 425
Qy      366 CTGGTGCCAGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACAT 425
Db      426 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 485
Qy      426 CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGCCAAGGGAGGTGTCCCT 485
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      486 NNNNNNNNNNNNNNNNCGAAGACATTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTT 545
Qy      486 GGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCGAGGTGGA 545
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      546 GGAACAGGAAGTCTTACTCCAGTGTGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605
Qy      546 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAACATCACGCG 605
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      606 NNNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTAT 665

```

Qy 606 GGAGCACAGCCTGGTGGCGACAGGCCGCTTGCTGACACGGCAACTACACCTGCGT 665
Db 666 TGATCACAAACCTCATATAAAGCAGGCCGACTCTGTGATACTGCAAATTATACCTGTGT 725

Qy 666 GGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTGTCATCGTCTACGTGAA 725
Db 726 TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACGTGCACTGTGTCATAGTCTATGTCAA 785

Qy 726 CGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCAGCGCCAGCTGGGCGCGGCTG 785
Db 786 CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGGACGAGGGTA 845

Qy 786 GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTCTGTGA 845
Db 846 TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCCTCTGTGA 905

Qy 846 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGCTGGAG 905
Db 906 AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGAC 965

Qy 906 CCCGTGGAGCAAGTGGTCGGCTGTGGCTGGACTGCACCCACTGGCGAGCCGTGAGTG 965
Db 966 GCCATGGAGCAAGTGGTCACTTGTGAACTGAGTGCACCCACTGGCGCAGGAGGGAGTG 1025

Qy 966 CTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGGCACTGACCTGGACACCCG 1025
Db 1026 CACGGGCCAGCCCCAAGAATGGAGGAAGGACTGCGATGCCCTCGTCTGCAATCAA 1085

Qy 1026 CAACTGTACCAGTGACCTCTGTGTACACAGTGCCTCTGGCCCTGAGGACGTGGCCCTCTA 1085
Db 1086 GAACTGCACTGATGGCTTGCAATGCAGACTGCTCTGATTAGATGATGTTGCTCTCTA 1145

Qy 1086 TGTGGG---CCTCATGCCGTGGCGTCTGCCCTGGCTGCTGCTGCTGTGCTCATCCT 1142
Db 1146 TGTTGGGATTGTGATAGCAGTGATCGTTGCCTGGCGATCTCTGTAGTTGCTGGCCTGTT 1205

Qy 1143 CGTTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC 1202
Db 1206 TGTGTATCGGAAGAATCATCGTGACTTGAGTCAGATATTATTGACTCTCGGCACTCAA 1265

Qy 1203 CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCATCTGCTCAC 1262
Db 1266 TGGGGCTTCAGCCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGCTGGC 1316

Qy 1263 CATCCAGCCGGACCTCAGCACCAACCACCCACCTACCAAGGGCAGTCTGTCCCCGGCA 1322
Db 1317 TGTACCCCCAGACCTCACGTGAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA 1376

Qy 1323 GGATG----GGCCCAGCCCCAAGTTCAGCTACCAAT----GGGCACCTGCTCAGCCC 1373
Db 1377 TGACGTCTCAGACAAAATCCAATGACCAACTCTCCAATTCTGGATCCACTGCCAACCT 1436

Qy 1374 CCTGGGTGGCGGCCACACACTGCACCCAGCTCTCCACCTCTGAGGCCGAGGAGTT 1433
Db 1437 GAAAATCAAAGTGTACAACACCTCAGGTGCTGTCAACCCCCAAGATGACCTCTGAGTT 1496

RESULT 7

AY411747

LOCUS AY411747 2532 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY411747
VERSION AY411747.1 GI:39767715
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES Location/Qualifiers
 source 1. .2532
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1. .>2532
 /locus_tag="HCM4327"
 ORIGIN
 Query Match 30.0%; Score 810.4; DB 9; Length 2532;
 Best Local Similarity 61.1%; Pred. No. 2e-162;
 Matches 1540; Conservative 0; Mismatches 851; Indels 129; Gaps 9;
 Qy 307 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGGGCTGGAGGAATAC 366
 |||||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTTTGGGGCTGGAGGATTAC 72
 Qy 367 TGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCCTACATC 426
 |||||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 73 TGGTGCCAGTGCCTGGAGCTCCGGGGCACCAAGAGTCGCCGAGCCTACGTC 132
 Qy 427 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCAAGGAGGTGTCCCTG 486
 |||||| |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGCAAGGAGGTGCCCTG 192
 Qy 487 GAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCGAGGTGGAG 546
 ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 193 GACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGTGCCTGTGCCGAGGTGGAA 252
 Qy 547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCCAATGTATACTACACCGGG 606
 |||||| | ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 253 TGGCTCAAGAACATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312
 Qy 607 GAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGCCAACTACACCTGCCGTG 666
 ||| ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 313 GACCAACAACCTCATCCGCCAGGCCGCCGTGCGACACTGCCAACTATACCTGCCGTG 372
 Qy 667 GCCAAGAACATCGTGCACGTGCCCGCAGGCCCTCCGCTGCTGTCATCGTCTACGTGAAC 726
 |||||| |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 373 GCCAAGAACATCGTGGCCAACGCCGGAGCACCAGGCCACTCGCCACCGTCATCGTCTACGTGAAT 432
 Qy 727 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGCGCTGG 786
 ||| ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 433 GGCGGCTGGTCCAGCTGGCAGAGTGGTACCCCTGCTCCAACCGCTGTGCCGAGGCTGG 492
 Qy 787 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCCCTCTCAACGGGGCGCTTCTGTGAG 846
 |||||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 493 CAGAAGCGCACCCGGACCTGCACCAACCCGCTCCACTCAACGGAGGGGCCCTGCGAG 552
 Qy 847 GGGCAGAACATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGCTGGAGC 906
 ||| ||| | ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db	553	GGCCAGGCATTCCAGAAGACCGCCTGCACCACCATCTGCCAGTCATGGGCGTGGACG	612
Qy	907	CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGAGCCGTGAGTGC	966
Db	613	GAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCACTGGCGTAGCCGAGTGC	672
Qy	967	TCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC	1026
Db	673	ATGGCGCCCCACCCAGAACCGGAGGCCGTGACTGCAGCGGACGCTGCTGACTCTAAG	732
Qy	1027	AACTGTACCAAGTACCTCTGTGTACACAGTGCTTCT-----	1062
Db	733	AACTGCACAGATGGCCTGTGCATGCAAAAATAAGAAAATCTAACAGCAGCCAC	792
Qy	1063	-----GGCCCTGAGGACGTGGCCCTCTATGTGGGCTC---ATGCCGTGGCCGTC	1110
Db	793	CTGCTGGAGGCCTCAGGGATGCGGCGCTGTATGCCGGCTCGTGGTGGCCATCTCGTG	852
Qy	1111	TGCCTGGCCTGCTGCTGCTGTCTCATCCTCGTTATTGCCGGAAGAAGGGAGGGCTG	1170
Db	853	GTCGTGGCAATCCTCATGGCGGTGGGGTGGTGTACCGCCGCAACTGCCGTGACTTC	912
Qy	1171	GACTCAGATGTGGCTGACTCGTCCATT---CTCACCTCAGGCTTCCAGCCCCTCAGC	1227
Db	913	GACACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTCCACCCGTCAACTT	972
Qy	1228	AAGCCCAGCAAAGCAGACAACCCCCATCTGCT-----CACCATCCAGCCGGACCTCAGC	1281
Db	973	AAGACGGCAAGGCCAGCAACCCGAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA	1032
Qy	1282	ACCACCAACCACCACTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG-----	1329
Db	1033	GCCAGGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAATC	1092
Qy	1330	-----CCCAGCCCCAAGTTCCAGCTCACCA--	1354
Db	1093	CCCATGACCAACTCTCCTCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC	1152
Qy	1355	-----ATGGGCACCTGCTCAGCCCCCTGGTGGCGGCC-----ACACA	1395
Db	1153	TCCAGCACACGGCTCTGGCCAGGCCTGGCAGATGGGCTGACCTGCTGGGGCTTG	1212
Qy	1396	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCGCCTCTCCACCCAG	1455
Db	1213	CCGCCTGGCACATACCCTAGCGATTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGC	1272
Qy	1456	AACTACTTCCGCTC-----CCTGCCCGAGGCCACAGCAACATGACCTAT	1500
Db	1273	GCCAGCCTCGGTTCCCAGCAGCTCTGGGCCTGCCCGAGACCCAGGGAGCAGCGTCAGC	1332
Qy	1501	GGGACCTTCAACTCCTCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCCTC	1560
Db	1333	GGCACCTTGGCTGCCTGGTGGAGGCTCAGCATCCCCGGCACAGGGTCAGCTGCTG	1392
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1393	GTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG	1452

Qy	2458 ATTTCCAGCCTGGACCCACCCTGTAGGCAGGGTGCCACTGGCGGACTCTGGCCCAGAAA	2517
Db	2293 TGCAACAGCCTAGATGCCCAACTCACGGGCAATGACTGGCGGATGTTAGCACAGAAAG	2352
Qy	2518 CTCCACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATC	2577
Db	2353 CTCTCTATGGACCGGTACCTGAATTACTTGCCACCAAAGCGAGCCCCACGGGTGTGATC	2412
Qy	2578 CTCAACCTGTGGAGGGCGCGGACTTCCCAACGGAACCTCAGCCAGCTGGCTGCAGCA	2637
Db	2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTAACAGCCTGGCGAGTGCC	2472
Qy	2638 GTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	2473 TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGGCCACCGACGGGACTGCTGA	2532

RESULT 8

AY411749

LOCUS AY411749 2532 bp DNA linear GSS 12-DEC-2003
 DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY411749
 VERSION AY411749.1 GI:39767717
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2532)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 2532)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES Location/Qualifiers
 source 1. .2532
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1. .>2532
 /locus_tag="HCM4327"
 ORIGIN

Query Match 28.9%; Score 780.4; DB 9; Length 2532;
 Best Local Similarity 60.1%; Pred. No. 5.1e-156;
 Matches 1515; Conservative 0; Mismatches 876; Indels 129; Gaps 8;

Qy	307	GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGCGAGAAGGTGTTGGGGCTGGAGGAATAC	366
Db	13	GAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACCTTCGGGCTCGAGGACTAC	72
Qy	367	TGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCCTACATC	426
Db	73	TGGTGCCAGTGCCTGGAGCTCTCGGGAACTACCAAGAGTCGCCAGCCTACATC	132
Qy	427	CGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCAAGGAGGTGTCCCTG	486
Db	133	CGCATTGCCTACTTGCAGAAGAACATTGACCAGGAGCCTCTGGCAAGGAGGTACCCCTG	192
Qy	487	GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCGAGGTGGAG	546
Db	193	GATCATGAGGTCCCTCTGCAGTGCGCCACCGGAGGGAGTGCCCTGTGGCTGAGGTGGAA	252
Qy	547	TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGACACCCCCAATGTATACTCACCGG	606
Db	253	TGGCTCAAGAACATGAAGATGTCATTGACCCGCTCAGGACACTAACTCCTGCTCACCA	312
Qy	607	GAGCACAGCCTGGTGGTGCAGACAGGCCCTTGCTGACACGCCAACTACACCTGCGTG	666
Db	313	GACCACAAACCTCATCCGCCAGGCGCGCTCTCAGACACGCCAACTACACCTGTGTG	372
Qy	667	GCCAAGAACATCGTGCACGTGCCCGCAGCGCCTCCGCTGCTGTCATCGTACGTGAAC	726
Db	373	GCCAAGAACATCGTGCACGCCAGCGCCGGAGCACCGCCACAGTCATCGTCTATGTGAAT	432
Qy	727	GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGCGCTGG	786
Db	433	GGAGGCTGGTCCAGCTGGCAGAGTGGTACCCCTGTTCAATCGCTGTGGCGAGGCTGG	492
Qy	787	CAGAACCGGAGCCGGAGCTGCACCAACCCGGCCCTCTAACGGGGCGCTTCTGTGAG	846
Db	493	CAGAACCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTCTGTGAG	552
Qy	847	GGGCAGAACATGTCCAGAAAACAGCCTGCGCACCCCTGTGCCAGTAGACGGCAGCTGGAGC	906
Db	553	GGACAGGCCCTCCAGAACAGACAGCTTGACCCACCGTGTGCCAGTGGATGGAGCGTGGACC	612
Qy	907	CCGTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGAGCCGTGAGTGC	966
Db	613	GAGTGGAGCAAGTGGTCTGCCCTGCAGCACAGAGTGTGCGCACTGGCGAGCCGCGAGTGC	672
Qy	967	TCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC	1026
Db	673	ATGGCACCGCCACCCAGAACGGAGGCCGTGACTGCAGCGGACGCTACTGACTCCAAG	732
Qy	1027	AACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC-----	1074
Db	733	AACTGCACGTGGCTGTGGCTGAATCAGAGAACACTCAAACGACCCCTAAAGCCAC	792
Qy	1075	-----GTGGCCCTCTATGTGGCCCTCATGCCGTGGCGTCTGC	1113

Db	1633	GAGGAGGTGGTACCTTGGATGAGGAGACCCCTAACACACCCCTGCTACTGCCAGCTGGAG 1692
Qy	1861	GCCAGTGCTGCTACGTCTCACCGAGCAGCTGGGCCGTTGCCCTGGTGGAGAGGCC 1920
Db	1693	GCTAAAGTCCTGCCACATCCTGCTGGACCAGCTGGTACCTACGTATTCATGGCGAGTCC 1752
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC 1980
Db	1753	TACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTCGCCCCAGCCCTGCACC 1812
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
Db	1813	TCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCTGTAGCACTGAAGGAG 1872
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC 2100
Db	1873	GTCCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAGCCAAAGCCTTGCTC 1932
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG 2160
Db	1933	TTTAAGGACAGTTACCACAAACCTACGCCCTCCCTCCATGACATCCCCATGCCACTGG 1992
Qy	2161	AAGAGTAAGCTCCTTGTAGCTACCAGGAGATCCCCTTTATCACATCTGGAATGGCACG 2220
Db	1993	AGGAGCAAACACTGGCCAAGTACCAGGAGATTCCCTTCTACCACGTCTGGAATGGCACG 2052
Qy	2221	CAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGTGACCTG 2280
Db	2053	CAGAGAGCCCTGCACCTTCACCCCTGGAGAGGCATAGCCTGGCCTCACGGAGTTC 2112
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Db	2113	ACCTGTAAGGTCTGCGTGCAGTCAGGGCAAGGGCAGATTTCCAGCTGCACACA 2172
Qy	2341	AACATCACC---AAGGACACAAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2397
Db	2173	ACGTTGGCCGAGACGCCCTGCTGGCTCCCTGGATGCTCTGCTCTGCCCGGGCAATGCC 2232
Qy	2398	GTCCCAGCCCTGGTGGGCCCAAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAGATA 2457
Db	2233	ATCACCAACCCAGCTGGACCCCTATGCCTCAAGATACCCCTGTCCATCGCCAAAAGATC 2292
Qy	2458	ATTTCCAGCCTGGACCCACCTGTAGGCAGGGTGCCACTGGCGACTCTGGCCCAGAAA 2517
Db	2293	TGCAGCAGCCTGGACGCCCAACTCCCAGGGCAACGACTGGAGGCTGTTGGCGCAGAAG 2352
Qy	2518	CTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGCCATGATC 2577
Db	2353	CTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGTCCCACAGGTGTATC 2412
Qy	2578	CTCAACCTGTGGAGGCGCGGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2637
Db	2413	TTAGACCTCTGGGAAGCTCGAACAGGATGACGGGACCTAACAGCCTGGCCAGTGCC 2472
Qy	2638	GTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2697
Db	2473	TTGGAGGAGATGGGCAAGAGTGAAGATGCTGGTAGCCATGGCACAGATGGCGATTGCTGA 2532

RESULT 9
BI758231
LOCUS BI758231 1034 bp mRNA linear EST 25-SEP-2001
DEFINITION 603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
mRNA sequence.
ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1034)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11501 row: g column: 20
High quality sequence stop: 793.
FEATURES Location/Qualifiers
source 1. .1034
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200171"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
ORIGIN
Query Match 28.5%; Score 768.6; DB 4; Length 1034;
Best Local Similarity 91.2%; Pred. No. 1.5e-153;
Matches 918; Conservative 0; Mismatches 74; Indels 15; Gaps 9;
QY 105 GCCTGGTGCCAACCGGACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACAT 164
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GCCTGGTGCCAACCGGACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACAT 60
Qy 165 CGTCAAGAACAAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCACGCAGATCTTCTT 224

Db	61	CGTCAAGAACAGGCCAGTGCTGCTTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTT	120
Qy	225	CAAGTGCACGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	284
Db	121	CAAGTGCACGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	180
Qy	285	GAGCAGTGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA	344
Db	181	GAGCAGTGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA	240
Qy	345	GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAC	404
Db	241	GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAC	300
Qy	405	CAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	464
Db	301	CAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTGCGCAAGAACTTCGAGCAGGAGCC	360
Qy	465	GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGG	524
Db	361	GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGG	420
Qy	525	CATCCCTCCAGCCAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	584
Db	421	CATCCCTCCAGCCAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	480
Qy	585	CCCCAATGTATAACATCACGCCGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGA	644
Db	481	CCCCAATGTATAACATCACGCCGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGA	540
Qy	645	CACGGCCAACCTACACCTGCGTGGCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGC	704
Db	541	CACGGACAACCTACACCTGCGTGGCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGC	600
Qy	705	TGCTGTACGTACGTGAAACGGTGGTGGTCGA-CGTGGACCGAGTGGTCCGTCTGCA	763
Db	601	TGCTGTACGTACGTGAAACGGTGGTGGTCGACCGTGGACCGAGTGGTCCGTCTGCA-	659
Qy	764	GCGCCAGCTGGCGCGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	823
Db	660	GCGCCAGCTGGCGCGCTGGCAGAACGGAGCCGGAGCTGCACAAACCCGGTGCCTC	719
Qy	824	TCAACGGGGCGCTTCTGTGA-GGGGCAGAACATGCCAGAAAACAGCCTGCCACCCCTG	882
Db	720	TCAACGGGGCGCTTCTGTGAGGGGCAGAACATGCCAGAACAGCAGC---TGCACCT	776
Qy	883	TGCCCAAGTAGACGGCAGCTGGAGCCGTGGAGCAAGTGGTGGCCTGT--GGGCTGGACT	940
Db	777	TGCCCAAGTAGACGGCAGCTGTAGCACGTGGAGCCAGTGGTGGCCTGTGGCTTGGCTT	836
Qy	941	GCACCCACT--GGCGGAGCCGTGAGTGCCTGACCCAGCACCCGCAACGGAGGGAGGA	998
Db	837	GCACCCACTGGCGAGCCGGAGTGCCTGTGACCCAGCACCCGGCACGGCAGGGAG	896
Qy	999	GTG---CCAGGGCACTGACCTGGACACCCGCAACTGTACCGAGTGCACCTGTGTACACAG	1055

Db 897 GTGTGCCAGGTCACTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCAC 954
 Qy 1056 TGCTTCTGCCCTGAGGACGTGCCCTCATGTGGCCTCATGCCG 1102
 | ||||||| ||||| ||||| ||||||| ||||| |||
 Db 955 GGGTTCTGCCCTGAGGACTTGCCTATGTGGCCTCATCCGG 1001

RESULT 10
 BC033727
 LOCUS BC033727 1532 bp mRNA linear HTC 25-MAR-2004
 DEFINITION Homo sapiens unc-5 homolog A (*C. elegans*), mRNA (cDNA clone IMAGE:5166762), containing frame-shift errors.
 ACCESSION BC033727
 VERSION BC033727.1 GI:21707230
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1532)
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1532)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghghi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 68 Row: i Column: 2

This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction. This clone has the following problem: frame shifted.

FEATURES source Location/Qualifiers
1. .1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166762"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

```

Query Match          27.3%; Score 736.6; DB 3; Length 1532;
Best Local Similarity 84.2%; Pred. No. 1.1e-146;
Matches 917; Conservative 0; Mismatches 4; Indels 168; Gaps 1;

Qy      292 GGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTT 351
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      612 GGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTT 671
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      352 GGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGT 411
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      672 GGGCTGGAGGAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGT 731
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      412 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACCTTCGAGCAGGAGCCGCTGG 471
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      732 CAGAAGGCCTACATCCGCATAGCCTATTGCGCAAGAACCTTCGAGCAGGAGCCGCTGG 791
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      472 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACC GGAGGGCATCC 531
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      792 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACC GGAGGGCATCC 851
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      532 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 591
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      852 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 911
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      592 GTATACATCACGCCGGAGCACAGCCTGGTGGCGACAGGCCGCCCTGCTGACACGGCC 651
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      912 GTATACATCACGCCGGAGCACAGCCTGGTGGCGACAGGCCGCCCTGCTGACACGGCC 971
        ||||||| | | | | | | | | | | | | | | | | | | | | | | | |

```

Qy	652	AACTACACCTGCGTGGCAAGAACATCGTGGCACGTGCCGAGCGCCTCCGCTGTC	711
Db	972	AACTACACCTGCGTGGCAAGAACATCGTGGCACGTGCCGAGCGCCTCCGCTGTC	1031
Qy	712	ATCGTCTACGTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGC	771
Db	1032	ATCGTCTACGTG-----	1043
Qy	772	TGTGGGCGGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCAGGCCCTCTAACAGGG	831
Db	1044	-----	1043
Qy	832	GGCGCTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCCACCCGTGCCAGTA	891
Db	1044	-----	1043
Qy	892	GACGGCAGCTGGAGCCGTGGAGCAAGTGGTCGGCTGTGGACTGCACCCACTGG	951
Db	1044	GACGGCAGCTGGAGCCGTGGAGCAAGTGGTCGGCTGTGGACTGCACCCACTGG	1103
Qy	952	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGTGCCAGGGACT	1011
Db	1104	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGTGCCAGGGACT	1163
Qy	1012	GACCTGGACACCCGCAACTGTACCATGACCTCTGTGTACACAGTGCTCTGCCCTGAG	1071
Db	1164	GACCTGGACACCCGCAACTGTACCATGACCTCTGTGTACACACTGCTCTGCCCTGAG	1223
Qy	1072	GACGTGGCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTT	1131
Db	1224	GACGTGGCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTT	1283
Qy	1132	GTCCTCATCCTCGTTATTGCCGGAAGAACGGAGGGCTGGACTCAGATGTGGCTGACTCG	1191
Db	1284	GTCCTCATCCTCGTTATTGCCGGAAGAACGGAGGGCTGGACTCAGATGTGGCTGACTCG	1343
Qy	1192	TCCATTCTCACCTCAGGCTTCCAGCCGTAGCATCAAGCCCAGCAAAGCAGACAACCCC	1251
Db	1344	TCCATTCTCACCTCAGGCTTCCAGCCGTAGCATCAAGCCCAGCAAAGCAGACAACCCC	1403
Qy	1252	CATCTGCTACCACCATCCAGCCGGACCTCAGCACCAACCAACCTACCAAGGGCAGTC	1311
Db	1404	CATCTGCTACCACCATCCAGCCGGACCTCAGCACCAACCAACCTACCAAGGGCAGTC	1463
Qy	1312	TGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGC	1371
Db	1464	TGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGC	1523
Qy	1372	CCCCTGGGT 1380	
Db	1524	CCCCTGGGT 1532	

RESULT 11
AI951556

LOCUS AI951556 788 bp mRNA linear EST 09-MAR-2000

DEFINITION wv36f04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:2531647 3' similar to TR:O08721 O08721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA sequence.
 ACCESSION AI951556
 VERSION AI951556.1 GI:5743866
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 788)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1125 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 446.
 FEATURES Location/Qualifiers
 source 1. .788
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2531647"
 /tissue_type="fibrotheoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Ov18"
 /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGACATTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 27.3%; Score 735.2; DB 1; Length 788;
 Best Local Similarity 96.5%; Pred. No. 2e-146;
 Matches 749; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 Qy 1424 CCGAGGAGTCGTCTCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCA 1483
 ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
 Db 1 CCGAGGAGTCGTCTCCGCCTCTCCGCCAGAACTACTTCCGCTCCCTGCCCGAGGCA 60
 Qy 1484 CCAGCAACATGACCTATGGGACCTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1543

Db	61	CCAGCAACATGACCTATGGACCTCAACTCCTCGGGGCCGCTGATGATCCATAA	120
Qy	1544	CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCT	1603
Db	121	CAGGAATCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCT	180
Qy	1604	ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGACCCCTGC	1663
Db	181	ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGACCCCTGC	240
Qy	1664	TGAGTCCCATCGTTAGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGG	1723
Db	241	TGAGTCCCATCGTTAGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGG	300
Qy	1724	CTATGGACCACTGTGGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCACT	1783
Db	301	CTATGGACCACTGTGGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCACT	360
Qy	1784	CGTGCAGGGCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCT	1843
Db	361	CGTGCAGGGCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCT	420
Qy	1844	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTTTG	1903
Db	421	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTATG	480
Qy	1904	CCCTGGTGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTG	1963
Db	481	CCCTGGTGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTG	540
Qy	1964	CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCC	2023
Db	541	CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATGACACTC	600
Qy	2024	ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGG	2083
Db	601	ACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGG	660
Qy	2084	AGCCACGGGTCTGCACTTCAAGGACAGTTACCAACACCTGCCCTATCCATCCACGATG	2143
Db	661	AGCCACGGTACTGCACTTCAAGGACAGTTACCAACACCTGCCCTATNCATCCACGATG	720
Qy	2144	TGCCCAAGCTCCCTGTGGAGAGTAAGCTCCTTGTCACTACCAGGAGATCCCTTT	2199
Db	721	TGCCCAAGCTNCCNTGTGGAGAGTAAGCTTCTGTCAGCTACCCAGAGATCCNCTAT	776

RESULT 12

BX348193/c

LOCUS BX348193 796 bp mRNA linear EST 08-APR-2004
 DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
 ACCESSION BX348193
 VERSION BX348193.2 GI:46286231
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30367258.
 Contact: Genoscope
 Genoscope - Centre National de Sequencing
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3239.r
 For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAF004ZD01_AF00293_1&c=3239.r

FEATURES
source Location/Qualifiers
 1. .796
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB008YE02"
 /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match	26.8%	Score	721.6	DB	5	Length	796
Best Local Similarity	96.4%	Pred.	No.	1.7e-143			
Matches	758	Conservative	0	Mismatches	26	Indels	2
Gaps	2						

Qy 1634 GGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTAGCTGTGGACCCCTG 1693
 ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 795 GTTGCCCCAAACCGGGCTGTCAGACCCCTGTTGAGTCCCATGGTAGCTGTGAA-CCCTG 737

Qy 1694 GCGTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCAGCCCTG 1753
 ||||||| ||||| | ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 736 GCGTCCTGCTCACCGGCCAGTCATCCTGGGTATGGACCACTGTGGGGAGCCAGCCCTG 677

Qy 1754 ACAGCTGGAGCCTGCGCCTC-AAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTG 1812
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 676 ACAGCTGAGGCCTGCGCCTCAAAAAAACAGTCGTGCGAGGACAGCTGGGAGTATGTCCTG 617

Qy 1813 CACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 1872
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 616 CACCTGGCGAGNAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 557

Qy 1873 TACGTCTTACCGAGCAGCTGGGCCGTTGCCCTGGTAGGGAGAGGCCCTCAGCGTGGCT 1932
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 556 TACGTCTTACCGAGCAGCTGGCCGCTTGCCCTGGTGGAGAGGCCCTAGCGTGGCT 497
 Qy 1933 GCCGCCAAGCGCCTCAAGCTGCCTGTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 1992
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 496 GCCGCCAAGCGCCTCAAGCTGCCTGTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 437
 Qy 1993 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 2052
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 436 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 377
 Qy 2053 GAGAACAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCACTTAAGGACAGT 2112
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 376 GAGAACAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCACTTAAGGACAGT 317
 Qy 2113 TACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTC 2172
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 316 TACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTC 257
 Qy 2173 CTTGTCAGCTACCAGGAGATCCCTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 2232
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 256 CTTGTCAGCTACCAGGAGATCCCTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 197
 Qy 2233 CACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 2292
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 196 CACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 137
 Qy 2293 TGGGTGTGGCAGGTGGAGGGCGACGGGAGGCTTCAGCATCAACTCAACATCACCAAG 2352
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 136 TGGGTGTGGCAGGTGGAGGGCGACGGGAGGCTTCAGCATCAACTCAACATCACCAAG 77
 Qy 2353 GACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAAGCGGGGGTCCCAGCCCTGGT 2412
 |||||||||||||||||||||||||||||||||||||||||||
 Db 76 GACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAAGCGGGGGTCCCAGNCCTGGT 17
 Qy 2413 GGCCCC 2418
 || |||
 Db 16 GGCCCC 11

RESULT 13

BI818609

LOCUS BI818609 818 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
 mRNA sequence.
 ACCESSION BI818609
 VERSION BI818609.1 GI:15929902
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 818)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11434 row: 1 column: 16
 High quality sequence stop: 744.

FEATURES	Location/Qualifiers
source	1. .818 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5174559" /lab_host="DH10B" /clone_lib="NIH_MGC_115" /note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match	26.5%	Score	713.4	DB	4	Length	818		
Best Local Similarity	97.6%	Pred. No.	9.4e-142						
Matches	745	Conservative	0	Mismatches	16	Indels	2	Gaps	2
Qy	1	ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC	60						
Db	4	ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC	63						
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120						
Db	64	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	123						
Qy	121	GACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA	180						
Db	124	GACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA	183						
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCACGGGAG	240						
Db	184	GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCACGGGAG	243						
Qy	241	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC	300						
Db	244	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC	303						
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	360						
Db	304	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	363						

Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGCACCACCAAGAGTCAGAAGGCC	420
Db	364	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGCACCACCAAGAGTCAGAAGGCC	423
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	424	TACATCCGCATAGCCTATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	483
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG	540
Db	484	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG	543
Qy	541	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	544	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	603
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCCAACATACACC	660
Db	604	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCCAACATACACC	663
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTC-GCCGCAGCGCCTCCGCTGCTGTCATCGTCTA	719
Db	664	TGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGCCCTCCGCTGCTGTCATCGTCTA	723
Qy	720	CGTGAACGGTGGTGGTCGACG-TGGACCGAGTGGTCCGTCTG	761
Db	724	CGTGGACGGCAGCTGGAGGCCGTTGGAGCCAGTGGTCGGGCTG	766

RESULT 14

AY411748

LOCUS	AY411748	2532 bp	DNA	linear	GSS	12-DEC-2003
DEFINITION	Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.					
ACCESSION	AY411748					
VERSION	AY411748.1 GI:39767716					
KEYWORDS	GSS.					
SOURCE	Pan troglodytes (chimpanzee)					
ORGANISM	Pan troglodytes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.					
REFERENCE	1 (bases 1 to 2532)					
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.					
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2 (bases 1 to 2532)					
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,					

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifier

source

1. .2532

/organis

/mol type="genomic DNA"

/db_xref="taxon:9598"

1-2533

gene

1. .22332
Locns. tag-WHGM4327

ORIGIN

Query Match 25.1%; Score 678.2; DB 9; Length 2532;
Best Local Similarity 54.4%; Pred. No. 3.7e-134;
Matches 1370; Conservative 0; Mismatches 1021; Indels 129; Gaps 9;

QY 307 GAGGTCCGCATTAATGTCCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 366

Qy 367 TGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCCTACATC 426

Db 73 TGGTGCCAGTGCCTGGCCTGGAGCTCTGCGGGCACCAAGAGTCGCCGAGCCTACGTC 132

Qy 547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCCAATGTATAACATCACCGGG 606

Qy 607 GAGCACAGCCTGGTGGTGCGACAGGCCGCTTGTGACACGGCCAATACACCTGCGTG 666

Db 313 GACCACAAACCTCATCATCCGCCAGGGCCGCTGTGCGGACACTGCCAACTATACCTGCGTG 372

Qy 667 GCCTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTACGTCTACGTGAAC 726
||||| ||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCCTGCCCCGTATCGTCTACGTGAAT 432

Db 433 GGC GGCTGGTCCAGCTGGGAGAGTGTCGCCCTGCTCCAACCGCTGTGGNNNGANNCTGG 492

Qy 787 CAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTCTGTGAG 846

947 CCCCCAGCATTCTGGCGAAGGACGGCTGGCCGACTAGACCCCCGCTGGAGG 952

Qy 907 CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC 966

Db	1453	GCAGAAAGTACCCCTCCGCTTCAGAAGGGACCCAGACAGTATTGAGCCCCCAGTGACC	1512
Qy	1681	TGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1513	TGTGGACCCACAGGCCTCCTGCTGTGCCACCCGTCATCCTCACCATGCCCACTGTGCC	1572
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1573	GAAGTCAGTGCCCGTGAUTGGATCTTCAGCTCAAGACCCAGGCCACCAGGGCCACTGG	1632
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1633	GAGGAGGTGGTGACCTGGATGAGGAGACCTGAACACACCCCTGCTACTGCCAGCTGGAG	1692
Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCGCTTGGCCCTGGTGGAGAGGCC	1920
Db	1693	CCCAGGGCCTGTACATCCTGCTGGACCAGCTGGCACCTACGTGTTACGGGCGAGTCC	1752
Qy	1921	CTCAGCGTGGCTGCCGCCAACGCCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Db	1753	TATTCCCNNNNNNCAGTCNNNNNGCTCCNGCTGGCCGTCTNGCCCTGCCCTCTGCACC	1812
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1813	TCCCTGGANNNNNNGCTCCGGGTCTACTGCCTGGNNNNNNNNNNNNNNNNNNNAG	1872
Qy	2041	GTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGCCTGCAC	2100
Db	1873	GTGCTGGAGCTGGAGCGGACTCTGGCGGATACTTGGTGGAGGAGCCAAACGCTAATG	1932
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	1933	TTCAAGGACAGTTACCACAAACCTGCGCTCCCTCCATGACCTCCCCATGCCATTGG	1992
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTATCACATCTGGAATGGCACG	2220
Db	1993	AGGAGCAAGCTGCTGCCAAATACCAGGAGATCCCCTCTGTACATTTGGAGTGGCAGC	2052
Qy	2221	CAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2053	CAGAAGGCCCTCCACTGCACCTCACCTGGAGAGGCACAGCTGGCCTCCACAGAGCTC	2112
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTC	2340
Db	2113	ACCTGCAAGATCTGCGTGCAGCAAGTGGAGGGAGGGCAGATATTCCAGCTGCATACC	2172
Qy	2341	AACATCACCAAG---GACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG	2397
Db	2173	ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACT	2232
Qy	2398	GTCcccAGCCCTGGTGGCCCCAGTGCCTTCAAGATCCCCTCCTCATCGCAGAAGATA	2457
Db	2233	GTCACCACCCAGCTGGACCTTATGCCTTCAAGATCCCAGTGTCCATCCGCCAGAAGATA	2292
Qy	2458	ATTTCCAGCCTGGACCCACCCCTGTAGGCGGGGTGCCACTGGGGACTCTGGCCCAGAAA	2517
Db	2293	TGCAACAGCCTAGATGCCCAACTCACGGGGCAATGACTGGGGATGTTAGCACAGAAG	2352

RESULT 15

BF311804
LOCUS BF311804 934 bp mRNA linear EST 21-NOV-2000
DEFINITION 601897316F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4126706 5',
mRNA sequence.
ACCESSION BF311804
VERSION BF311804.1 GI:11259566
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 934)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1016 row: p column: 03
High quality sequence stop: 707.
FEATURES
source Location/Qualifiers
1. .934
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4126706"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 24.9%; Score 672.6; DB 2; Length 934;
 Best Local Similarity 93.3%; Pred. No. 5.1e-133;
 Matches 747; Conservative 0; Mismatches 49; Indels 5; Gaps 4;

Qy	1071 GGACGTGGCCCTCATGGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTGCT 1130
Db	2 GGACGTGGCCCTCATGGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTGCT 61
Qy	1131 TGTCCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTC 1190
Db	62 TGTCCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTC 121
Qy	1191 GTCCATTCTCACCTCAGGCTTCAGCCCCGTCAAGCCCAGCAAAGCAGACAACCC 1250
Db	122 GTCCATTCTCACCTCAGGCTTCAGCCCCGTCAAGCCCAGCAAAGCAGACAACCC 181
Qy	1251 CCATCTGCTCACCATCCAGCCGACCTCAGCACCAACCACCTACCAGGGCAGTCT 1310
Db	182 CCATCTGCTCACCATCCAGCCGACCTCAGCACCAACCACCTACCAGGGCAGTCT 241
Qy	1311 CTGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAG 1370
Db	242 CTGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAG 301
Qy	1371 CCCCCGGTGGCGGCCACACACTGCACCACTGCACAGCTCTCCACCTCTGAGGCCAGGA 1430
Db	302 CCCCCGGTGGCGGCCACACACTGCACCACTGCACAGCTCTCCACCTCTGAGGCCAGGA 361
Qy	1431 GTTCGTCTCCGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAA 1490
Db	362 GTTCGTCTCCGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAA 421
Qy	1491 CATGACCTATGGGACCTTCAACTTCCGCCCCGGCTGATGATCCCTAACAGGTAT 1550
Db	422 CATGACCTATGGGACCTTCAACTTCCGCCCCGGCTGATGATCCCTAACAGGAAT 481
Qy	1551 CAGCCTCCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCAC 1610
Db	482 CAG-CTCCTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCAC 540
Qy	1611 GCTGCACAAGCCGGAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCC 1670
Db	541 GCTGCACAAGCCGGAGACGTGAGGTTG-CCCTAGCTGGCTGTCAGACCCCTGCTGAGTCC 599
Qy	1671 CATCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCC--TGGCTATG 1728
Db	600 CATCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCCTGGCTAATG 659
Qy	1729 GACCACTGTGGGAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGTCGTGC 1788
Db	660 GACCCACTGTGGGAAGCCAGCCTGACAGTTGGAGCCTGGCTCAAAAAGGAGTCGGTGC 719
Qy	1789 GAGGGCAGCTGGGAGGATGTGCTGACCTGGCGAGGGAGGCCCTCCACCTACTAC 1848
Db	720 GAGGGCAGCTGGGA-GATGTGCTGGACTGGGGAAAGAAGGGGGCCTCCACCCCTATTAT 778

Qy 1849 TGCCAGCTGGAGGCCAGTGCC 1869
 ||| |||| |||||||| |||||
Db 779 TGGCAGTGGGAGGCAGGCC 799

Search completed: March 6, 2005, 10:10:35
Job time : 8489.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:11:26 ; Search time 11456.2 Seconds
(without alignments)
11407.261 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggccggcct.....tgtcgaggctgagtgtga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query				Description
		Match	Length	DB	ID	
1	2697	100.0	2752	6	AX449572	Sequence
2	2687.4	99.6	2697	6	AX451652	Sequence
3	2621.4	97.2	2881	6	AX527916	Sequence

4	2435.4	90.3	2784	6	CQ730306	CQ730306 Sequence
5	2343	86.9	3580	6	AX367094	AX367094 Sequence
6	2297	85.2	3992	10	MMU487852	AJ487852 Mus muscu
7	2252.2	83.5	2697	6	AX268596	AX268596 Sequence
8	2252.2	83.5	2697	10	RNU87305	U87305 Rattus norv
9	2252.2	83.5	3014	6	BD057524	BD057524 Netrin re
10	1957.4	72.6	3844	10	BC058084	BC058084 Mus muscu
11	1625.4	60.3	2688	9	BC009333	BC009333 Homo sapi
12	1552.4	57.6	1787	6	BD057525	BD057525 Netrin re
13	1302.8	48.3	9700	6	AX054976	AX054976 Sequence
14	992	36.8	9299	10	MMU72634	U72634 Mus musculu
15	988.8	36.7	2962	5	AY187310	AY187310 Gallus ga
16	986	36.6	9328	10	AB118026	AB118026 Rattus no
17	970	36.0	3646	6	CQ881060	CQ881060 Sequence
18	970	36.0	3646	9	AF055634	AF055634 Homo sapi
19	942.4	34.9	3770	9	AY126437	AY126437 Homo sapi
20	936.2	34.7	2860	6	AX686445	AX686445 Sequence
21	936.2	34.7	2860	6	AX686447	AX686447 Sequence
22	912.8	33.8	3672	10	BC048162	BC048162 Mus muscu
23	912.8	33.8	3672	10	BC057560	BC057560 Mus muscu
24	912.2	33.8	2895	6	AX512281	AX512281 Sequence
25	904.2	33.5	2986	6	CQ881052	CQ881052 Sequence
26	904.2	33.5	3933	6	CQ850929	CQ850929 Sequence
27	904.2	33.5	3933	9	AK128132	AK128132 Homo sapi
28	904.2	33.5	3935	9	AB096256	AB096256 Homo sapi
29	901	33.4	3884	6	AR528525	AR528525 Sequence
30	901	33.4	3884	6	AX464012	AX464012 Sequence
31	901	33.4	3884	9	AY358351	AY358351 Homo sapi
32	892.4	33.1	2995	6	AX497288	AX497288 Sequence
33	892	33.1	4294	10	AK122575	AK122575 Mus muscu
34	874.6	32.4	3788	10	MMU487853	AJ487853 Mus muscu
35	852.4	31.6	2838	10	RNU87306	U87306 Rattus norv
36	843.6	31.3	2625	6	CQ721377	CQ721377 Sequence
37	841.4	31.2	2831	6	BD057526	BD057526 Netrin re
38	764.8	28.4	813	6	AX054892	AX054892 Sequence
39	751.4	27.9	4330	5	AY744919	AY744919 Petromyzo
40	738.4	27.4	2230	6	CQ845766	CQ845766 Sequence
41	738.4	27.4	2230	9	AK131380	AK131380 Homo sapi
42	717.2	26.6	2832	5	AY099459	AY099459 Xenopus l
43	668.4	24.8	2612	6	CQ881064	CQ881064 Sequence
44	623.2	23.1	2661	6	AX800717	AX800717 Sequence
45	623.2	23.1	2868	6	AX800719	AX800719 Sequence

ALIGNMENTS

RESULT 1

AX449572

LOCUS AX449572 2752 bp DNA linear PAT 03-JUL-2002
 DEFINITION Sequence 1 from Patent WO0210216.
 ACCESSION AX449572
 VERSION AX449572.1 GI:21698195
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS 1 Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
 Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
 Macdougall,J.R. and Smithson,G.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;
 Curagen Corporation (US)
 FEATURES Location/Qualifiers
 source 1. .2752
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 2697; DB 6; Length 2752;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 105
 Qy 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
 Qy 121 GACCTGCTCCCCACTTCCTGGTGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 166 GACCTGCTCCCCACTTCCTGGTGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 225
 Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGAG 240
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 226 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGAG 285
 Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
 |||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 286 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 345
 Qy 301 ACCATGGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
 |||||||||||||||||||||||||||||||||||||||||||||||
 Db 346 ACCATGGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 405
 Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCCACCAAGAGTCAGAAGGCC 420
 |||||||||||||||||||||||||||||||||||||||||||||||
 Db 406 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCCACCAAGAGTCAGAAGGCC 465
 Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 480
 |||||||||||||||||||||||||||||||||||||||||||
 Db 466 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 525
 Qy 481 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
 |||||||||||||||||||||||||||||||||||||||||||
 Db 526 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585
 Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAAATGTATACATC 600
 |||||||||||||||||||||||||||||||||||||||||||

Db	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCAATGTATACTC	645
Qy	601	ACGCAGGGAGCACAGCCTGGTGGCGACAGGCCGCTTGCTGACACGGCAAACACACC	660
Db	646	ACGCAGGGAGCACAGCCTGGTGGCGACAGGCCGCTTGCTGACACGGCAAACACACC	705
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTACATCGTCTAC	720
Db	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTACATCGTCTAC	765
Qy	721	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC	780
Db	766	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC	825
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	840
Db	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	885
Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	900
Db	886	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	945
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT	960
Db	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT	1005
Qy	961	GAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC	1020
Db	1006	GAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC	1065
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1125
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGTCTGCTGCTGCTGTGCTCATC	1140
Db	1126	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGTCTGCTGCTGCTGTGCTCATC	1185
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1186	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAAGGGCAGTCTCTGTCCCCGG	1320
Db	1306	ACCATCCAGCCGGACCTCAGCACCAACCACCAACCTACCAAGGGCAGTCTCTGTCCCCGG	1365
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGT	1380
Db	1366	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGT	1425
Qy	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCC	1440
Db	1426	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCC	1485

Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1486	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1545
Qy	1501	GGGACCTTCAACTTCTCGGGGGCCGGTGTGATGATCCCTAATACAGGTATCAGCCTCTC	1560
Db	1546	GGGACCTTCAACTTCTCGGGGGCCGGTGTGATGATCCCTAATACAGGTATCAGCCTCTC	1605
Qy	1561	ATCCCCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1606	ATCCCCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCCATCGTTAGC	1680
Db	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCCATCGTTAGC	1725
Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1726	TGTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1785
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1846	GAGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Qy	1861	GCCAGTGCCTGCTACGTCTTACCGGAGCAGCTGGCGCGCTTGCCCTGGTGGAGAGGCC	1920
Db	1906	GCCAGTGCCTGCTACGTCTTACCGGAGCAGCTGGCGCGCTTGCCCTGGTGGAGAGGCC	1965
Qy	1921	CTCAGCGTGGCTGCCGCCAACGCGCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Db	1966	CTCAGCGTGGCTGCCGCCAACGCGCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	2025
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2100
Db	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2145
Qy	2101	TTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2146	TTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2220
Db	2206	AAGAGTAAGCTCCTTGTCACTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2265
Qy	2221	CAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2266	CAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2325

Qy 2281 GCCTGCAAGCTGTGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTC 2340
 |||||||
 Db 2326 GCCTGCAAGCTGTGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTC 2385
 |||||||
 Qy 2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2400
 |||||||
 Db 2386 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2445
 |||||||
 Qy 2401 CCAGCCCTGGTGGGCCAGTGCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT 2460
 |||||||
 Db 2446 CCAGCCCTGGTGGGCCAGTGCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT 2505
 |||||||
 Qy 2461 TCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
 |||||||
 Db 2506 TCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565
 |||||||
 Qy 2521 CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
 |||||||
 Db 2566 CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
 |||||||
 Qy 2581 AACCTGTGGAGGCGCGGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
 |||||||
 Db 2626 AACCTGTGGAGGCGCGGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685
 |||||||
 Qy 2641 GCTGGACTGGGCCAGCCAGCCAGCTGGCTCTTCACAGTGTGGAGGCTGAGTGCTGA 2697
 |||||||
 Db 2686 GCTGGACTGGGCCAGCCAGACGCTGGCTCTTCACAGTGTGGAGGCTGAGTGCTGA 2742

RESULT 2

AX451652

LOCUS AX451652 2697 bp DNA linear PAT 03-JUL-2002
 DEFINITION Sequence 1 from Patent WO0233080.
 ACCESSION AX451652
 VERSION AX451652.1 GI:21698587
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Koehler, R.H.
 TITLE Regulation of human netrin binding membrane receptor unc5h-1
 JOURNAL Patent: WO 0233080-A 1 25-APR-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES Location/Qualifiers
 source 1. .2697
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 2687.4; DB 6; Length 2697;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGGCCGTCCGGCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGTTGGCTC 60

Db |||||||
1 ATGGCCGTCCGCCGCGCTGTGGCCAGCAGTCCTGGCATAGTCCTGCCGCTGGCTC 60

Qy 61 CGCGGCTCGGTGCCAGCAGAGTGCACCGTGGCAACCCAGTGCCTGGTGCACCCG 120
|||
Db 61 CGCGGCTCGGTGCCAGCAGAGTGCACCGTGGCAACCCAGTGCCTGGTGCACCCG 120

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAGCCA 180
|||
Db 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAGCCA 180

Qy 181 GTGCTGCTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGAG 240
|||
Db 181 GTGCTGCTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGAG 240

Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGCTGCC 300
|||
Db 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGCTGCC 300

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
|||
Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGCACCAAGAGTCAGAACGCC 420
|||
Db 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGCACCAAGAGTCAGAACGCC 420

Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCCTGGCAAGGAGGTG 480
|||
Db 421 TACATCCGCATAGCCTATTCGCAAGAACTTCGAGCAGGAGCCCTGGCAAGGAGGTG 480

Qy 481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCAG 540
|||
Db 481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCAG 540

Qy 541 GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC 600
|||
Db 541 GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC 600

Qy 601 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCGCTTGCTGACACGGCAACTACACC 660
|||
Db 601 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCGCTTGCTGACACGGCAACTACACC 660

Qy 661 TGCCTGGCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGCTGCTGTATCGTCTAC 720
|||
Db 661 TGCCTGGCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGCTGCTGTATCGTCTAC 720

Qy 721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 780
|||
Db 721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 780

Qy 781 GGCTGGAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGTTTC 840
|||
Db 781 GGCTGGAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGTTTC 840

Qy 841 TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCGTGCCCAGTAGACGGCAGC 900
|||

Db	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCAGTGGACGGCAGC	900
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGCCCTGAGGACGTGGCC	1080
Qy	1081	CTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTTGTCCCTCATC	1140
Db	1081	CTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTTGTCCCTCATC	1140
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTGTCCCCGG	1320
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Qy	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTCTCC	1440
Db	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTCTCC	1440
Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Qy	1501	GGGACCTTCAACTCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCTC	1560
Db	1501	GGGACCTTCAACTCCTGGGGCCGGCTGATGATCCCTAACAGGAATCAGCCTCTC	1560
Qy	1561	ATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1681	TGTGGACCCCTGGCTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740

Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCAGCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCAGCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGAGAGGCC	1920
Db	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGAGAGGCC	1920
Qy	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Db	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCACGATGTGCCAGCTCCCTGTGG	2160
Db	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCACGATGTGCCAGCTCCCTGTGG	2160
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAACCTGCGCTATCCATCACGATGTGCCAGCTCCCTGTGG	2220
Db	2161	AAGAGTAAGCTCCTTGTCACTACCAACCTGCGCTATCCATCACGATGTGCCAGCTCCCTGTGG	2220
Qy	2221	CAGCGGTACTTGCACCTTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2221	CAGCGGTACTTGCACCTTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAAGCTTCAGCATCAACTTC	2340
Qy	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Qy	2401	CCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCCCTCCTCATTGGCAGAAGATAATT	2460
Db	2401	CCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCCCTCCTCATTGGCAGAAGATAATT	2460
Qy	2461	TCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAGAAACTC	2520
Db	2461	TCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAGAAACTC	2520
Qy	2521	CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db	2521	CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580

Qy	2581 AACCTGTGGGAGGCAGCAGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG . 2640
Db	2581 AACCTGTGGGAGGCAGCAGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
Qy	2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTTCACAGTGTGGAGGCTGAGTGCTGA 2697
Db	2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTTCACAGTGTGGAGGCTGAGTGCTGA 2697

RESULT 3

AX527916

LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002
 DEFINITION Sequence 1 from Patent WO0229038.
 ACCESSION AX527916
 VERSION AX527916.1 GI:25172359
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Herrmann, J.L., Rastelli, L. and Shimkets, R.A.
 TITLE Novel proteins and nucleic acids encoding same and antibodies
 directed against these proteins
 JOURNAL Patent: WO 0229038-A 1 11-APR-2002;
 Curagen Corporation (US)
 FEATURES Location/Qualifiers
 source 1. .2881
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 2621.4; DB 6; Length 2881;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 2673; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

Qy	1 ATGGCCGTCCGGCCCGGCCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 60
Db	87 ATGGCCGTCCGGCCCGGCCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 146
Qy	61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db	147 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Qy	121 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
Db	207 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 266
Qy	181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACGGGGAG 240
Db	267 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACGGGGAG 326
Qy	241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	327 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy	301	ACCATGGAGGTCCGCATTAATGTCCTAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCCTAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCTGGACCCGTCCTGGACCCAAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCTGGACCCGTCCTGGACCCAAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGCCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGCCAACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCGCTGCTGTATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCGCTGCTGTATCGTCTAC	806
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGGCCTCTCAACGGGGCGCTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGGCCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCCACCTGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTTGCTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCCTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCCTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCACGGAGGGGAGGAGTGCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCACGGAGGGGAGGAGTGCAGGGCACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGCCTCATGCCGTGCCGTCTGCCCTGGTCTGCTGCTGCTTGTCCCTC	1137
Db	1167	GCCCTCTATGTGGCCTCATGCCGTGCCGTCTGCCCTGGTCTGCTGCTGCTTGTCCCTC	1226
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197

Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCAACCACCTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTCGTC	1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTCAACTTCCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTCAACTTCCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCTACGTCTCACCGAGCAGCTGGCCGCTTGGCTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCTACGTCTCACCGAGCAGCTGGCCGCTTGGCTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTACTGCCTGCATGACACCCACGATGCACTC	2034

Db 2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
 Qy 2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTC 2094
 |||||
 Db 2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTC 2183
 Qy 2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2154
 |||||
 Db 2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2243
 Qy 2155 CTGTGGAAGAGTAAGCTCCTTGTCAAGCTACCAGGAGATCCCCTTTATCACATCTGGAAT 2214
 |||||
 Db 2244 CTGTGGAAGAGTAAGCTCCTTGTCAAGCTACCAGGAGATCCCCTTTATCACATCTGGAAT 2303
 Qy 2215 GGCACGCAGCGGTACTTGCACTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT 2274
 |||||
 Db 2304 GGCACGCAGCGGTACTTGCACTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT 2363
 Qy 2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
 |||||
 Db 2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
 Qy 2335 AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
 |||||
 Db 2424 AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
 Qy 2395 GGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCATTGGCAGAACAG 2454
 |||||
 Db 2484 GGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCATTGGCAGAACAG 2543
 Qy 2455 ATAATTCCAGCCTGGACCCACCTGTAGGCGGGTGCCACTGGCGACTCTGGCCAG 2514
 |||||
 Db 2544 ATAATTCCAGCCTGGACCCACCTGTAGGCGGGTGCCACTGGCGACTCTGGCCAG 2603
 Qy 2515 AAACCTCACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
 |||||
 Db 2604 AAACCTCACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
 Qy 2575 ATCCTCAACCTGTGGAGGCAGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
 |||||
 Db 2664 ATCCTCAACCTGTGGAGGCAGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
 Qy 2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2694
 |||||
 Db 2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2783
 Qy 2695 TGA 2697
 |||
 Db 2784 TGA 2786

RESULT 4
CQ730306

LOCUS CQ730306 2784 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 16240 from Patent WO02068579.
 ACCESSION CQ730306
 VERSION CQ730306.1 GI:42303801

KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16240 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .2784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.3%; Score 2435.4; DB 6; Length 2784;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 6; Indels 177; Gaps 3;
Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Db 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAAACCCAGTGCCTGGTGCACCCG 120
Db 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAAACCCAGTGCCTGGTGCACCCG 120
Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
Db 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 240
Db 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 240
Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTCGGCTGGAG 360
Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTCGGCTGGAG 360
Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 420
Db 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 420
Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db 421 TACATCCGCATAGCCATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Qy 481 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Db	481	TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATACTC	600
Db	541		
Db	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATACTC	600
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCCAACCTACACC	660
Db	601		
Db	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCCAACCTACACC	660
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	720
Db	661		
Db	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	720
Qy	721	-----	720
Db	721	GGTGGGCCCGGGACTCCCTGGTCACAGGGAGAGGCAGTCGGTGCCCCGGCAGTGAC	780
Qy	721	-----GTGAACGGTGGGTGGTCACGTGGACCGAGTGG	753
Db	781		
Db	781	ATGTGGCTGTCCCTCTGTCCGGCCAGTGAACGGTGGGTGGTCACGTGGACCGAGTGG	840
Qy	754	TCCGTCTGCAGGCCAGCTGTGGCGCGCTGGCAGAAACGGAGGCCGGAGCTGCACCAAC	813
Db	841		
Db	841	TCCGTCTGCAGGCCAGCTGTGGCGCGCTGGCAGAAACGGAGGCCGGAGCTGCACCAAC	900
Qy	814	CCGGCGCCTCTCAACGGGGCGCTTCTGTGAGGGGAGAATGTCCAGAAAACAGCCTGC	873
Db	901		
Db	901	CCGGCGCCTCTCAACGGGGCGCTTCTGTGAGGGGAGAATGTCCAGAAAACAGCCTGC	960
Qy	874	GCCACCTGTGCCAGTAGACGGCAGCTGGAGGCCGTGGAGCAAGTGGTGGCCTGTGGG	933
Db	961		
Db	961	GCCACCTGTGCCAGTAGACGGCAGCTGGAGGCCGTGGAGCAAGTGGTGGCCTGTGGG	1020
Qy	934	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGG	993
Db	1021		
Db	1021	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGG	1080
Qy	994	GAGGAGTGCCAGGGACTGACCTGGACACCCGCAACTGTACCAAGTGACCTCTGTGTACAC	1053
Db	1081		
Db	1081	GAGGAGTGCCAGGGACTGACCTGGACACCCGCAACTGTACCAAGTGACCTCTGTGTACAC	1140
Qy	1054	A-----GTGCTTCTGGCCCT	1068
Db	1141		
Db	1141	AACTCCTACACCCCTGCCCAACCAAGGCCATGCTGTCTCCCGAGCTGCTTCTGGCCCT	1200
Qy	1069	GAGGACGTGGCCCTCTATGTGGCCTCATGCCGTGCCGTCTGCCTGGCCTGCTGCTG	1128
Db	1201		
Db	1201	GAGGACGTGGCCCTCTATGTGGCCTCATGCCGTGCCGTCTGCCTGGCCTGCTGCTG	1260
Qy	1129	CTTGTCTCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGAC	1188
Db	1261		
Db	1261	CTTGTCTCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGAC	1320
Qy	1189	TCGTCCATTCTCACCTCAGGCTCCAGGCCGTAGCATCAAGCCCAGCAAAGCAGACAAC	1248
Db	1321		
Db	1321	TCGTCCATTCTCACCTCAGGCTCCAGGCCGTAGCATCAAGCCCAGCAAAGCAGACAAC	1380

Qy	1249	CCCCATCTGCTACCACATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGT	1308
Db	1381	CCCCATCTGCTACCACATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGT	1440
Qy	1309	CTCTGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTC	1368
Db	1441	CTCTGTCCCCGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTC	1500
Qy	1369	AGCCCCCTGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTTGAGGCCAG	1428
Db	1501	AGCCCCCTGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTTGAGGCCAG	1560
Qy	1429	GAGTCGCTCCGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGC	1488
Db	1561	GAGTCGCTCCGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGC	1620
Qy	1489	AACATGACCTATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGT	1548
Db	1621	AACATGACCTATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGA	1680
Qy	1549	ATCAGCCTCCTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTC	1608
Db	1681	ATCAGCCTCCTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTC	1740
Qy	1609	ACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGT	1668
Db	1741	ACGCTGCACAAGCCGGAAGACGTG-----	1764
Qy	1669	CCCATCGTTAGCTGTGGACCCCTGGCGTCTGCTCACCCGGCAGTCATCCTGGCTATG	1728
Db	1765	-----AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCAGTCATCCTGGCTATG	1815
Qy	1729	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTG	1788
Db	1816	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTG	1875
Qy	1789	GAGGGCAGCTGGGAGGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTAC	1848
Db	1876	GAGGGCAGCTGGGAGGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTAC	1935
Qy	1849	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTG	1908
Db	1936	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTG	1995
Qy	1909	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTGCCCG	1968
Db	1996	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTGCCCG	2055
Qy	1969	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGAT	2028
Db	2056	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGAT	2115
Qy	2029	GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCA	2088
Db	2116	GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCA	2175

Qy	2089	CGGGTCCTGCACCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCC	2148
Db	2176	CGGGTCCTGCACCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCC	2235
Qy	2149	AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCACTGACCTTCACCCCTGGAGCGTGT	2208
Db	2236	AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCACTGACCTTCACCCCTGGAGCGTGT	2295
Qy	2209	TGGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCCTGGAGCGTGT	2268
Db	2296	TGGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCCTGGAGCGTGT	2355
Qy	2269	ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCT	2328
Db	2356	ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCT	2415
Qy	2329	AGCATCAACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGA	2388
Db	2416	AGCATCAACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGA	2475
Qy	2389	GAAGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTTCATT	2448
Db	2476	GAAGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTTCATT	2535
Qy	2449	CAGAACGATAATTCCAGCCTGGACCCACCCGTAGGCGGGTGCCACTGGCGACT	2508
Db	2536	CAGAACGATAATTCCAGCCTGGACCCACCCGTAGGCGGGTGCCACTGGCGACT	2595
Qy	2509	GCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCC	2568
Db	2596	GCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCC	2655
Qy	2569	GCCATGATCCTCAACCTGTGGAGGCAGCGCAGCTTCCCAACGGCAACCTCAGCC	2628
Db	2656	GCCATGATCCTCAACCTGTGGAGGCAGCGCAGCTTCCCAACGGCAACCTCAGCC	2715
Qy	2629	GCTGCAGCAGTGGCTGGACTGGCCAGCCAGCCTGGCTCTTCACAGTGTGGAGG	2688
Db	2716	GCTGCAGCAGTGGCTGGACTGGCCAGCCAGCCTGGCTCTTCACAGTGTGGAGG	2775
Qy	2689	GAGTGCTGA 2697	
Db	2776	GAGTGCTGA 2784	

RESULT 5

AX367094

LOCUS	AX367094	3580 bp	DNA	linear	PAT 16-FEB-2002
DEFINITION	Sequence 13 from Patent WO0198354.				
ACCESSION	AX367094				
VERSION	AX367094.1 GI:18855296				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				

AUTHORS Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
 Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
 Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
 Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
 Sanjanwala,M.S.
TITLE Receptors
JOURNAL Patent: WO 0198354-A 13 27-DEC-2001;
 Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
 source 1. .3580
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 6052371CB1"
ORIGIN
 Query Match 86.9%; Score 2343; DB 6; Length 3580;
 Best Local Similarity 93.6%; Pred. No. 0;
 Matches 2524; Conservative 0; Mismatches 5; Indels 168; Gaps 1;
 Qy 1 ATGGCCGTCGGGCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGCTTGGCTC 60
 Db 4 ATGGCCGTCGGGCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGCTTGGCTC 63
 Qy 61 CGCGGCTCGGGTGCCACAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCACACCCG 120
 Db 64 CGCGGCTCGGGTGCCACAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCACACCCG 123
 Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
 Db 124 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA 183
 Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCACGGGGAG 240
 Db 184 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCACGGGGAG 243
 Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
 Db 244 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 303
 Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
 Db 304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 363
 Qy 361 GAATACTGGTGCAGTGCAGTGGCATGGAGCTCCTGGGCACCAAGAGTCAGAACGCC 420
 Db 364 GAATACTGGTGCAGTGCAGTGGCATGGAGCTCCTGGGCACCAAGAGTCAGAACGCC 423
 Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 480
 Db 424 TACATCCGCATAGCCTATTCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 483
 Qy 481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
 Db 484 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543
 Qy 541 GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCAAATGTATACATC 600

Db	1216	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTTGAGGCCGAGGAGTTCGTCTCC	1275
Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1276	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1335
Qy	1501	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1336	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1395
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1396	ATCCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1455
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCCATCGTTAGC	1680
Db	1456	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCCATCGTTAGC	1515
Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1516	TGTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1575
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1576	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1635
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1636	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1695
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1920
Db	1696	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1755
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC	1980
Db	1756	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC	1815
Qy	1981	TCCCTCGAGTACAACATCCGGCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1816	TCCCTCGAGTACAACATCCGGCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAG	1875
Qy	2041	GTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC	2100
Db	1876	GTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC	1935
Qy	2101	TTCAAGGACAGTTACCAACCTGCCCTATCCATCCACGATGTGCCAGCTCCGTGG	2160
Db	1936	TTCAAGGACAGTTACCAACCTGCCCTATCCATCCACGATGTGCCAGCTCCGTGG	1995
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAACCTGCCCTATCCATCCACGATGTGCCAGCTCCGTGG	2220
Db	1996	AAGAGTAAGCTCCTTGTCACTACCAACCTGCCCTATCCATCCACGATGTGCCAGCTCCGTGG	2055
Qy	2221	CAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2056	CAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2115

Qy 2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
 |||||||
 Db 2116 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2175
 |||||||
 Qy 2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2400
 |||||||
 Db 2176 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2235
 |||||||
 Qy 2401 CCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCCTCCTCATTGGCAGAAGATAATT 2460
 |||||||
 Db 2236 CCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCCTCCTCATTGGCAGAAGATAATT 2295
 |||||||
 Qy 2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGACTCTGGCCCAGAAACTC 2520
 |||||||
 Db 2296 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGACTCTGGCCCAGAAACTC 2355
 |||||||
 Qy 2521 CACCTGGACAGCCATCTCAGCTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC 2580
 |||||||
 Db 2356 CACCTGGACAGCCATCTCAGCTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC 2415
 |||||||
 Qy 2581 AACCTGTGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
 |||||||
 Db 2416 AACCTGTGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2475
 |||||||
 Qy 2641 GCTGGACTGGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGCTGA 2697
 |||||||
 Db 2476 GCTGGACTGGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGCTGA 2532

RESULT 6

MMU487852

LOCUS MMU487852 3992 bp mRNA linear ROD 24-SEP-2002
 DEFINITION Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).
 ACCESSION AJ487852
 VERSION AJ487852.1 GI:22035783
 KEYWORDS netrin receptor Unc5h1; Unc5h1 gene.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Engelkamp,D.
 TITLE Cloning of three mouse Unc5 genes and their expression patterns at
 mid-gestation
 JOURNAL Mech. Dev. 118 (1-2), 191-197 (2002)
 MEDLINE 22239710
 PUBMED 12351186
 REFERENCE 2 (bases 1 to 3992)
 AUTHORS Engelkamp,D.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for
 Brain Research, Deutscherstrasse 46, Frankfurt 60528, GERMANY
 FEATURES Location/Qualifiers
 source 1. .3992
 /organism="Mus musculus"
 /mol_type="mRNA"

gene /db_xref="taxon:10090"
1. .3992
/gene="Unc5h1"
CDS 232. .2928
/gene="Unc5h1"
/codon_start=1
/product="netrin receptor Unc5h1"
/protein_id="CAD32250.1"
/db_xref="GI:22035784"
/db_xref="GOA:Q8K1S4"
/db_xref="UniProt/TrEMBL:Q8K1S4"
/translation="MAVRPGLWPALLGIVLTAWLRSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVCKAVPATQIFFKCNGEWWRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKFQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSDLPNVYITREHSLVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEERCAGADLDTRNCTSDDLCLHTSSGPEDVALYIGLVAVAVCLLLLLVLVLIYCRKKEGLDSDVADSSIITSGFQPVSIKPSKADNPPLLTIQPDLSSTTTTYQGSILCPRQDGSPPKFQLSNGHILLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSMAYGTFNFLGGRLMIPNTGISLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEESPShLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLFAPVACTSLEYNIRVYCLHDTHDALKEVVALEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSKLVSYQEIPFYHIWNGTQQYLHCTFTLERVNASTSDLACKVWWQVEGDQSFNINFNITKDTRFAEMLALESEGGVPAVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPGNGNLGQLAAAVAGLQPDAGLFTVSEAEC"

ORIGIN

```

Query Match           85.2%; Score 2297; DB 10; Length 3992;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 2447; Conservative 0; Mismatches 250; Indels 0; Gaps 0;

    1 ATGGCCGTCCGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGCTTGGCTC 60
       ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
232 ATGGCCGTCCGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTT 291

    61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
292 CGTGGTTGGGTGCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCC 351

    121 GACCTGCTCCCCACTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAGCCA 180
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
352 GACCTGCTGCCCACTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAGCCC 411

    181 GTGCTGCTTGTGTGCAAGGCCGTGCCACGCAGATCTTCTTAAGTGCAACGGGAG 240
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
412 GTGCTGCTGGTGTGCAAGGCTGTGCCACCCAGATCTTCTTAAGTGCAACGGGAA 471

    241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGCTGCC 300
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
472 TGGGTTGCCAGGTGATCACGTATCGAACGCAGCACTGACGGCAGCAGCGGATTGCCA 531

    301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
532 ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 591

```

Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	592	GAGTACTGGTGCCAGTGTGTCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC	651
Qy	421	TACATCCGCATAGCCAGATTGCAGAAGAACCTCGAGCAGGAGGCCCTGGCCAAGGAGGTG	480
Db	652	TACATCCGGATTGCCTATTGCGCAAGAACCTTGAGCAGGAGGCCCTGGCCAAGGAAGTG	711
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	712	TCACTGGAGCAAGGCATTGTGCTACCTTGTGCCCCCGGAAGGAATCCCCCAGCTGAG	771
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	772	GTGGAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTGACCCCAATGTGTACATC	831
Qy	601	ACGCGGGAGCACAGCCTGGTGGCGACAGGCCGCCTGCTGACACGGCCAACATACACC	660
Db	832	ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCGCCTGGCCGACACTGCCAACTACACC	891
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	720
Db	892	TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGGCCCTGCGGCCGTATTGTTAT	951
Qy	721	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	952	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGCGT	1011
Qy	781	GGCTGGCAGAACGGAGGCCGGAGCTGCACCAACCGGCCCTCTAACGGGGCGCTTTC	840
Db	1012	GGCTGGCAGAACGGAGGCCGGAGCTGCACCAACCGGCCACCTCTAACGGGGCGCTTC	1071
Qy	841	TGTGAGGGGCAGAACATGTCCAGAAAACAGCCTGCGCCACCCACTGTGCCAGCAGC	900
Db	1072	TGTGAGGGGCAGAACATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCAGTGGATGGAGC	1131
Qy	901	TGGAGCCCGTGGAGCAAGTGGTGGCCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT	960
Db	1132	TGGAGCCCAGGAGTAAGTGGTCAGCCTGCCGCTTGACTGCACCCACTGGCGAGCCGG	1191
Qy	961	GAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCACGGGACTGACCTGGAC	1020
Db	1192	GAGTGCTCCGACCCAGCGCCCCGACGGAGGTGAGGAGTGCCTGGGTGCTGACCTGGAC	1251
Qy	1021	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1252	ACCCGCAACTGTACCAAGTGACCTCTGCCCTGCACACCTCTCCGCCCGAGGACGTGGCT	1311
Qy	1081	CTCTATGTGGGCCTCATGCCGTGGCCGTCTGCCCTGGCTGCTGCTGCTGTGCTCATC	1140
Db	1312	CTCTACATCGGCCCTCGCCGTGGCCGTGCTCATCTTGCTGCTGCTGGTCCTCGTC	1371
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1372	CTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATCCTT	1431
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260

Db	1432	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGTC	1491
Qy	1261	ACCATCCAGCCGGACCTCAGCACCACCACCTACCAGGGCAGTCTGTCCCCGG	1320
Db	1492	ACCATCCAACCGGACCTCAGCACCACCACGACCACCTACCAGGGCAGCCTGTGTCCCCGG	1551
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1552	CAGGATGGACCCAGCCCCAAGTCCAGCTCTCTAAATGGCACCTGCTCAGCCCACACTGGGC	1611
Qy	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1612	AGTGGCCGCCATACGCTGCACCACAGCTCCCCACCTTGAGGCTGAGGACTTCGTCTCC	1671
Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1672	CGCCTCTCACCCAAAACTACTTCCGTTCTGCCCGCGGTACCAGAACATGGCCTAT	1731
Qy	1501	GGGACCTTCAACTTCCCTCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCCTC	1560
Db	1732	GGGACCTTCAACTTCCCTCGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTC	1791
Qy	1561	ATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1792	ATACCCCCGGACGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAG	1851
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCATCGTTAGC	1680
Db	1852	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCATCGTTAGC	1911
Qy	1681	TGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGGCTATGGACCAGTGTGGG	1740
Db	1912	TGTGGGCCAGGAGTCCTGCTCACCGGCCAGTCATCCTGCCATGGACCAGTGTGGG	1971
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1972	GAGCCCAGTCCCACAGCTGGAGCCTGCCCTCAAAAGCAGTCCTGTGAGGGCAGCTGG	2031
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	2032	GAGGACGTGCTGCACCTGGTGAGGAGTCGCCCTCTCATCTACTACTGCCAGCTGGAG	2091
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGGCCCTTGCCTGGTGGAGAGGCC	1920
Db	2092	GCGGGGCCCTGCTATGTCTTCACCGAGCAGCTAGGCCGCTTGCCTGGTGGAGAGGCC	2151
Qy	1921	CTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC	1980
Db	2152	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTGTGTTGCCCTGTGGCTGTACG	2211
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	2212	TCCCTCGAGTACAACATCCGAGTGTACTGCCCTGCACGACACCCACGATGCTCTCAAGGAG	2271
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCGCAC	2100

Db 2272 GTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGCAC 2331
 Qy 2101 TTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCCGTGG 2160
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2332 TTCAAAGACAGTTACCACAACCTACGTCTGCCATCCACGACGTGCCAGCTCCGTGG 2391
 Qy 2161 AAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAATGGCAG 2220
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2392 AAGAGCAAGCTCCTTGTCACTACAGGAGATCCCCTTTACCACATCTGGAATGGCACT 2451
 Qy 2221 CAGCGGTACTTGCACCTCACCCCTGGAGCGTGTAGCCCCAGCACTAGTGACCTG 2280
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2452 CAGCAGTATCTGCACCTCACCCCTGGAGCGCGTCAATGCCAGCACCAGGACCTG 2511
 Qy 2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACTC 2340
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2512 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTT 2571
 Qy 2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2400
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2572 AACATCACTAAGGACACGAGGTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGTC 2631
 Qy 2401 CCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCCTCATTGGCAGAAGATAATT 2460
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2632 CCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCCTCATTGGCAAAAGATCATT 2691
 Qy 2461 TCCAGCCTGGACCCACCCGTAGGCGGGTGCCGACTGGCGACTCTGGCCAGAAACTC 2520
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2692 ACCAGCCTGGACCCACCCGTGAGCCGGCGACTGGCGAAGTCTAGCCCAGAAACTT 2751
 Qy 2521 CACCTGGACAGCCATTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC 2580
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2752 CACCTGGACAGCCATCTAGCTTCTTGCCCTCAAGCCCAGCCCTACAGCCATGATCCTC 2811
 Qy 2581 AACCTGTGGAGGCAGCGGCACCTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 2812 AACCTATGGGAGGCAGCGGCACCTCCCCAACGGCAACCTCGGCCAGCTGGCCAGCTGTG 2871
 Qy 2641 GCTGGACTGGGCCAGCCAGACGCTGGCTCTTACAGTGTGGAGGCTGAGTGCTGA 2697
 ||| ||||| ||||| ||||| ||||| |||||
 Db 2872 GCCGGACTGGGCCAGCCAGATGCTGGCTCTCACCGTGTAGAGGCCAGTGCTGA 2928

RESULT 7

AX268596

LOCUS	AX268596	2697 bp	DNA	linear	PAT 29-OCT-2001
DEFINITION	Sequence 15 from Patent WO0175440.				
ACCESSION	AX268596				
VERSION	AX268596.1 GI:16541710				
KEYWORDS					
SOURCE	Rattus sp.				
ORGANISM	Rattus sp. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1				
AUTHORS	Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and				

Pratt, J.Y.
 TITLE Schizophrenia related genes
 JOURNAL Patent: WO 0175440-A 15 11-OCT-2001;
 WELFIDE CORPORATION (JP)
 FEATURES Location/Qualifiers
 source 1. .2697
 /organism="Rattus sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:10118"
 ORIGIN
 Query Match 83.5%; Score 2252.2; DB 6; Length 2697;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;
 Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCAGCGCTCCTGGCATAGTCCTGCCGCTTGGCTC 60
 |||||||
 Db 1 ATGGCCGTCCGGCCCGGCCTGTGGCAGTGCTCCTGGCATAGTCCTGCCGCTTGGCTC 60
 |||||
 Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCACACCG 120
 |||||
 Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCCACGGTGGCAATCCAGTGCACACCG 120
 |||||
 Qy 121 GACCTGCTTCCCCACTTCCTGGAGGCCGAGGATGTGTACATCGTAAGAACAGCCA 180
 |||||||
 Db 121 GACCTGCTGCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAGCCA 180
 |||||
 Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCACGCAGATCTTCAAGTGCACAGGGAG 240
 |||||
 Db 181 GTGTTGTGGTGTGCAAGGCCGTGCCACGCAGATCTTCAAGTGCACAGGGAG 240
 |||||
 Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGCTGCC 300
 |||||
 Db 241 TGGGTGCCAGGTGATCACGTAATTGAAACGCAGCACCGACAGCAGCAGGGATTGCCA 300
 |||||
 Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTCGGCTGGAG 360
 |||||
 Db 301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTGTCGGCTGGAG 360
 |||||
 Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGCACCACCAAGAGTCAGAACAGGCC 420
 |||||
 Db 361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTGGTACCAACCAAGTCAGAACAGGCC 420
 |||||
 Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGTGCCAAGGAGGTG 480
 |||||
 Db 421 TACATCCGGATTGCCATTGCGCAAGAACCTTGAGCAGGAGCCACTGCCAAGGAAGTG 480
 |||||
 Qy 481 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
 |||||
 Db 481 TCACTGGAGCAAGGCATTGTACTACCTTGTGCCCTCCAGAAGGAATCCCCCAGCTGAG 540
 |||||
 Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
 |||||
 Db 541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTGATCCAAATGTGTACATC 600
 |||||
 Qy 601 ACGCGGGAGCACAGCCTGGTGGTGCAGAGGCCGCTTGTGACACGCCAACTACACC 660

Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCGCTGGCCACACGGCCAATACACC	660
Qy	661	TGC GTGGCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTCATCGTCTAC	720
Db	661	TGTGTGGCAAGAACATCGTAGCCGTCGCCGAAGCACCTCTGCAGCGGTATTGTTAT	720
Qy	721	GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC	780
Db	721	GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGT	780
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGCTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTAACGGGGCGCTTC	840
Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCGTGCCCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCCAGTGGATGGAGC	900
Qy	901	TGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTCTGTGGAGTAAGTGGTCAGCCTGTGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGCTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1021	ACCCGCAACTGTACCAAGTGACCTCTGCCCTGCACACCGCTTCTGCCCGAGGACGTGGCT	1080
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTGTGCTCATC	1140
Db	1081	CTCTACATCGGCCCTTGTGCTGTGGCTGTGCTCTTCTGCTGTTGCTGGCCCTTGGGA	1140
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTTACTGTCGAAGAAGGAAGGGCTGGACTCCGATGTGGCGACTCGTCCATCCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCAGACCTCAGCACCAACCACCACTACCACCTACCAGGGCAGTCTATGTTGAGG	1320
Qy	1321	CAGGATGGCCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGACCCAGCCCCAAGTCCAGCTCTAATGGCACCTGCTCAGCCCCACTGGGG	1380
Qy	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGGCAGGAGTCGTCTCC	1440
Db	1381	AGTGGCCGCCATACGTTGCACCAAGCTCACCCACCTCTGAGGGTGGAGACTTCGTCTCC	1440
Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1441	CGCCTCTCACCCAAAACACTTTGTTCCCTGCCCGCGGCACCAGCAACATGGCCTAC	1500

Qy	1501	GGGACCTTCAACTCCTCGGGGGCCGGTGTGATGCCATAACAGGTATCAGCCTCCTC	1560
Db	1501	GGGACCTTCAACTCCTCGGGGGCCGGTGTGATGCCATAACAGGGATCAGCCTCCTC	1560
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATACCCCCGGATGCCATCCCCCGAGGAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCGTCCGTCAACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1681	TGTGGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTGCAATGGACCACTGTGGA	1740
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACCTGGTGAGGAGTCACCTTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCTGGTAGGAGAGGCC	1920
Db	1861	GCCGGGGCCTGCTATGTCTCACGGAGCAGCTGGCCGCTTGCCTGGTAGGAGAGGCC	1920
Qy	1921	CTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTGTGTTGCGCCGGTGGCCTGCACC	1980
Db	1921	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTCTGTTGCTCCGTGGCCTGTACG	1980
Qy	1981	TCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGGCCACGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2101	TTCAAAGACAGTTACCACAAACCTACGTCTCTCCATCCACGACGTGCCAGCTCCCTGTGG	2160
Qy	2161	AAGAGTAAGCTCCTTGTCACTTACCAACGAGATCCCTTTATCACATCTGGAATGGCACG	2220
Db	2161	AAGAGCAAGCTACTTGTCACTTACCAACGAGATCCCTTTACACATCTGGAACGGCACC	2220
Qy	2221	CAGCGGTACTTGCACCTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTACCTG	2280
Db	2221	CAGCAGTATCTGCACCTCACCCCTGGAGCGCATCAACGCCAGCACCAGCACCTG	2280
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2281	GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGCAGAGCTTCACATCAACTTC	2340

Qy	2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2341 AACATCACTAAGGACACAAGGTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGTC	2400
Qy	2401 CCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2460
Db	2401 CCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2460
Qy	2461 TCCAGCCTGGACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
Db	2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT	2520
Qy	2521 CACCTGGACAGCCATCTCAGCTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db	2521 CACCTGGACAGCCATCTTAGCTCTTGCCCTCAAGCCCAGCCCTACAGCCATGATCCTC	2580
Qy	2581 AACCTGTGGAGGCAGCAGCTTCCCAACGGAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Db	2581 AACCTATGGGAGGCACGGCACTTCCCAACGGAACCTCGGCCAGCTGGCAGCAGCTGTG	2640
Qy	2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	2641 GCCGGACTGGCCAACCAGATGCTGGCCTTTCACGGTGTGGAGGCCAGTGTTGA	2697

RESULT 8

RNU87305

LOCUS	RNU87305	2697 bp	mRNA	linear	ROD	15-MAY-1997
DEFINITION	Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.					
ACCESSION	U87305					
VERSION	U87305.1 GI:2055391					
KEYWORDS						
SOURCE	Rattus norvegicus (Norway rat)					
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 2697)					
AUTHORS	Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L. and Tessier-Lavigne, M.					
TITLE	Vertebrate homologues of <i>C. elegans</i> UNC-5 are candidate netrin receptors					
JOURNAL	Nature 386 (6627), 833-838 (1997)					
MEDLINE	97271897					
PUBMED	9126742					
REFERENCE	2 (bases 1 to 2697)					
AUTHORS	Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K. and Tessier-Lavigne, M.					
TITLE	Direct Submission					
JOURNAL	Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San Francisco, CA 94143-0452, USA					
FEATURES	Location/Qualifiers					
source	1. .2697 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /tissue_type="brain and ventral spinal cord"					

CDS /dev_stage="18 day embryo and 13 day embryo"
1. .2697
/codon_start=1
/product="transmembrane receptor UNC5H1"
/protein_id="AAB57678.1"
/db_xref="GI:2055392"
/translation="MAVRPGLWPVLLGIVLAAWRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLILVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSGLPTMEVRINVSQQVEKVGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSDLPNVYITREHSILVVRQARLADTANYTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVKTACATLCPVVDGSWSSWKSACGLDCTHWRSRECSDPAPRNGGEERCRGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLALGLIYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLTLIQPDLSTTTTYQGSILCSRQDGSPKFQLSNGHILSPLGSGRHTLHHSSPTSEAEDFVSRSLTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLIIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEESPShLYYCQLEAGACYVFTEQLGRFALVGEALSAATKRLRLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVALEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWQVEGDGQSFNINFNITKDTRFAELLAESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARHFPGNGNLGQLAAAVAGLQPDAGLTVSEAEc"

ORIGIN

Query Match 83.5%; Score 2252.2; DB 10; Length 2697;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

Qy	1	ATGGCCGTCCGGCCCGGCCTGTGCCAGCGCTCCTGGCATAGTCCTGCCGCTGGCTC	60
Db	1	ATGGCCGTCCGGCCCGGCCTGTGCCAGCGCTCCTGGCATAGTCCTGCCGCTGGCTT	60
Qy	61	CGCGGCTCGGGTGCCAGCAGACTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG	120
Db	61	CGTGGTTGGGTGCCAGCAGACTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCC	120
Qy	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAGCCA	180
Db	121	GACCTGCTGCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAGCCG	180
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGAG	240
Db	181	GTGTTGGTGGTGTGCAAGGCTGTGCCACCCAGATCTTCTTCAAGTGCAATGGGAA	240
Qy	241	TGGGTGCCAGGTGGACCACGTGATCGAGCGCACGACAGACGGGAGCAGTGGCTGCC	300
Db	241	TGGGTCCGCCAGGTCATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA	300
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	360
Db	301	ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAAGTGTGTTGGCTGGAG	360
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	361	GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCAACCAAAGTCAGAAGGCC	420
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGCCAAGGAGGTG	480

Db	421	TACATCCGGATTGCCTATTGCGCAAGAACCTTGAGCAGGAGGCCACTGCCAAGGAAGTG	480
Qy	481	TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	481	TCACTGGAGCAAGGCATTGTACTACCTGTGCCCCCCAGAAGGAATCCCCCAGCTGAG	540
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACTC	600
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCAATGTGTACATC	600
Qy	601	ACGCAGGGAGCACAGCCTGGTGGTGCACAGGCCCTTGCTGACACGCCAATACACC	660
Db	601	ACGCAGGGAGCACAGCCTAGTCGTGCGTCAGGCCCTGGCCACACGCCAATACACC	660
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCCCTGCCGAAGCACCTCTGCAGCGGTATTGTTAT	720
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	721	GTGAACGGTGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGT	780
Qy	781	GGCTGGCAGAAACGGAGGCCGGAGCTGCACCAACCCGGCCCTCTAACGGGGCGCTTC	840
Db	781	GGCTGGCAGAAACGGAGGCCGGAGCTGCACCAACCCGGCACCTCTAACGGGGCGCTTC	840
Qy	841	TGTGAGGGGCAGAACATGTCCAGAAAACAGCCTGCCACCCCTGTGCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGGCAGAACATGTCCAGAAAACAGCCTGCCACCTCTGTGCCAGTGGATGGAGC	900
Qy	901	TGGAGCCCCTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGAGCCGT	960
Db	901	TGGAGTCGTGGAGTAAGTGGTCAGCCTGTGGCTTGACTGCACCCACTGGCGAGCCGC	960
Qy	961	GAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAAGTGACCTCTGCCCTGCACACCGCTTCTGCCCGAGGACGTGGCT	1080
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGTCTGCTGCTGCTTGTCCCTCATC	1140
Db	1081	CTCTACATCGGCCTTGTGCGTGTGGCTGTGTGCCCTTCTGCTGTTGTGCTGGCCCTTGG	1140
Qy	1141	CTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTACTGTGCCAAGAAGGAAGGGCTGGACTCCGATGTGGCGACTCGTCCATCCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCGTACGCATCAAGCCCAGCAAAGCAGACAACCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCACCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTGTCCCCGG	1320

Db	1261	ACCATCCAGCCAGACCTCAGCACCACTACCACCTACCAGGGCAGTCTATGTTGAGG	1320
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGACCCAGCCCCAAGTCCAGCTCTTAATGGCACCTGCTCAGCCCAGTGGG	1380
Qy	1381	GCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTCCTC	1440
Db	1381	AGTGGCCGCCATACGTTGCACCACAGCTACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1441	CGCCTCTCACCCAAAACTACTTCGTTCCCTGCCCGCGCACAGAACATGGCCTAC	1500
Qy	1501	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCCTC	1560
Db	1501	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAACAGGGATCAGCCTCCTC	1560
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1681	TGTGGGCCAGGAGTCCTGCTCACCGGCCAGTCATCCTTGAATGGACCACTGTGGA	1740
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCCCTCAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACCTGGTGAGGAGTCACCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCGCTTGCCTGGTGGAGAGGCC	1920
Db	1861	GCCGGGGCCTGCTATGTCTCACGGAGCAGCTGGCGCTTGCCTGGTAGGAGAGGCC	1920
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGCAC	1980
Db	1921	CTCAGCGTGGCTGCCACCAAGGCCCTCAGGCTCTTGTGCTCCGTGGCCTGTACG	1980
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCATACGACACCCACGACGCTCTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	2100
Qy	2101	TTCAAAGACAGTTACCAACCTGCGCTATCCATCACGATGTGCCAGCTCCGTGAG	2160
Db	2101	TTCAAAGACAGTTACCAACCTACGTCTCCATCCACGACGTGCCAGCTCCGTGAG	2160

Qy 2161 AAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAATGGCACG 2220
 ||||| ||||| ||||||| ||||||| ||||| ||||| ||||| |||||
 Db 2161 AAGAGCAAGCTACTTGTCACTACCAAGGAGATCCCTTTACCACATCTGGAACGGCAC 2220
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Qy 2221 CAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG 2280
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Db 2221 CAGCAGTATCTGCACCTTCACCCCTGGAGCGCATCAACGCCAGCACCAGGACCTG 2280
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Qy 2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACTTC 2340
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Db 2281 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGCAGAGCTTCAACATCAACTTC 2340
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Qy 2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2400
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Db 2341 AACATCACTAAGGACACAAGGTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGTC 2400
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Qy 2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTCCTCATTGGCAGAAGATAATT 2460
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Db 2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTCCTCATTGGCAGAAGATAATT 2460
 ||||| ||||| ||||||| ||||||| ||||| ||||| |||||
 Qy 2461 TCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCCAGAAACTC 2520
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Db 2461 GCCAGTCTGGACCCACCCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Qy 2521 CACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC 2580
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Db 2521 CACCTGGACAGCCATCTTAGCTTCTTGCCCTCAAGCCCAGCCCTACAGCCATGATCCTC 2580
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Qy 2581 AACCTGTGGAGGCAGCGGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Db 2581 AACCTATGGGAGGCAGGGCACTTCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Qy 2641 GCTGGACTGGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGCTGA 2697
 ||||| ||||| ||||||| ||||| ||||| ||||| |||||
 Db 2641 GCCGGACTGGCCAACCAGATGCTGGCTTTCACGGTGTGGAGGCCAGTGTTGA 2697

RESULT 9

BD057524

LOCUS	BD057524	3014 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Netrin receptors.				
ACCESSION	BD057524				
VERSION	BD057524.1 GI:22603130				
KEYWORDS	JP 2001505062-A/1.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct other sequences; artificial sequences.				
REFERENCE	1 (bases 1 to 3014)				
AUTHORS	Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.				
TITLE	Netrin receptors				
JOURNAL	Patent: JP 2001505062-A 1 17-APR-2001; THE REGENTS OF THE UNIV OF CALIFORNIA				
COMMENT	PN JP 2001505062-A/1 PD 17-APR-2001 PF 19-FEB-1998 JP 1998536840 PR 19-FEB-1997 US 08/808982				

PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
 PI MASU,
 PI KAZUKO KEINO MASU
 PC C07K1/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC
 Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
 source 1..3014
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 83.5%; Score 2252.2; DB 6; Length 3014;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

Qy	1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 60
Db	1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTGCCGCTGGCTT 60
Qy	61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db	61 CGTGGTTGGGTGCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120
Qy	121 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
Db	121 GACCTGCTGCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAAGCCG 180
Qy	181 GTGCTGTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGAG 240
Db	181 GTGTTGTTGGTGTGCAAGGCTGTGCCACCCAGATCTTCAAGTGCAATGGGAA 240
Qy	241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	241 TGGGTCCGCCAGGTGATCACGTAATTGAACGCAGCACCGACAGCAGCGGATTGCCA 300
Qy	301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db	301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTGGCTGGAG 360
Qy	361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAACGCC 420
Db	361 GAATACTGGTGCCTGGCATGGAGCTCCTGGGTACCAACCAAGAGTCAGAACGCC 420
Qy	421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db	421 TACATCCGGATTGCCTATTCGCGCAAGAACTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480
Qy	481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db	481 TCACTGGAGCAAGGCATTGTACTACCTTGTGCCACCCAGAAGGAATCCCCCAGCTGAG 540
Qy	541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db	541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy	601	ACGCAGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCAACTACACC	660
Db	601	ACGCAGGGAGCACAGCCTAGTCGTGCGTCAGGCCGCCTGGCCACACGGCAACTACACC	660
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCGTGCCGAAGCACCTCTGCAGCGTCATTGTTAT	720
Qy	721	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC	780
Db	721	GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGCTGCAGGCCAGCTGTGGCGT	780
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCCTCTCACACGGGGCGCTTC	840
Db	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCACACGGGGCGCTTC	840
Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCACCCACTGTGCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCACCTGTGCCAGTGGATGGAGC	900
Qy	901	TGGAGCCCCTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT	960
Db	901	TGGAGTCGTGGAGTAAGTGGTCAGCCTGTGGCTTGACTGCACCCACTGGCGAGCCGC	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGCTGGGTGCTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCCTGCACACCGCTTCTGCCCGAGGACGTGGCT	1080
Qy	1081	CTCTATGTGGGCCTCATGCCGTGGCCGTCTGCCCTGGCTGCTGCTGCTGTGCCTCATC	1140
Db	1081	CTCTACATCGGCCCTTGTGCGCTGTGGCTGTGCTCTTCTTGCTGTTGCTGGCCCTTGG	1140
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCATTCTC	1200
Db	1141	CTCATTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCGACTCGTCATCCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAAGGGCAGTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCAGACCTCAGCACCAACCACCTACCAAGGGCAGTCTATGTTGAGG	1320
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGACCCAGCCCCAAGTCCAGCTCTAATGGCACCTGCTCAGCCCAGTGGG	1380
Qy	1381	GGCGGCCGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCC	1440
Db	1381	AGTGGCCGCCATACGTTGCACCAAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440

Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1441	CGCCTCTCCACCCAAA ACTACTTCGTTCCCTGCCCGGGACCAGAACATGGCCTAC	1500
Qy	1501	GGGACCTTCAACTCCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1501	GGGACCTTCAACTCCTCGGGGCCGGCTGATGATCCCTAATACGGGATCAGCCTCCTC	1560
Qy	1561	ATCCCCCAGATGCCATACCCCAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATACCCCCGGATGCCATACCCCAGGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCAGTCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1681	TGTGGGCCAGGAGTCCTGCTCACCCGCCAGTCATCCTTGAATGGACCCTGTGGG	1740
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCCCTCAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qy	1801	GAGGATGTGCTGCACTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACTTGGTGAGGAGTCACCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1861	GCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCCTGGGGAGAGGCC	1920
Db	1861	GCCGGGGCCTGCTATGTCTTACGGAGCAGCTGGCCGCTTGCCTGGTAGGAGAGGCC	1920
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Db	1921	CTCAGCGTGGCTGCCACCAAGGCCCTCAGGCTCCTGTGCTCCGTGGCCTGTACG	1980
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2101	TTCAAAGACAGTTACCACAAACCTACGTCTCCATCCACGACGTGCCAGCTCCCTGTGG	2160
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAATGGCACG	2220
Db	2161	AAGAGCAAGCTACTTGTCACTACCAAGGAGATCCCTTTACACATCTGGAACGGCACG	2220
Qy	2221	CAGCGGTACTTGCACCTTCACCCGGAGCGTGTGAGCCAGCACTAGTGACCTG	2280
Db	2221	CAGCAGTATCTGCACCTTCACCCGGAGCGCATCAACGCCAGCACCAGGACCTG	2280
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340

Db	2281	GCCTGCAAGGTGTGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTAACATCAACCTTC 2340
Qy	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC 2400
Db	2341	
Db	2341	AACATCACTAAGGACACAAGGTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGTC 2400
Qy	2401	CCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCCTCCTCATTCGGCAGAAGATAATT 2460
Db	2401	
Db	2401	CCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCCTCCTCATTCGGCAGAAGATAATT 2460
Qy	2461	TCCAGCCTGGACCCACCCGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db	2461	
Db	2461	GCCAGTCTGGACCCACCCGTGCAGCCGGGCGCGACTGGAGAACTCTAGCCCAGAAACTT 2520
Qy	2521	CACCTGGACAGCCATTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
Db	2521	
Db	2521	CACCTGGACAGCCATTTAGCTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580
Qy	2581	AACCTGTGGAGGCAGCGCCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
Db	2581	
Db	2581	AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCTGTG 2640
Qy	2641	GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA 2697
Db	2641	
Db	2641	GCCGGACTGGCCAACCAGATGCTGGCCTTTCACGGTGTGGAGGCCAGTGTTGA 2697

RESULT 10

BC058084

LOCUS BC058084 3844 bp mRNA linear ROD 21-OCT-2003
 DEFINITION Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone MGC:66671 IMAGE:6813463), complete cds.
 ACCESSION BC058084
 VERSION BC058084.1 GI:34784158
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3844)
 REFERENCE AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 3844)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (08-SEP-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 126 Row: b Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 23346570.

FEATURES	Location/Qualifiers
source	1. .3844 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="MGC:66671 IMAGE:6813463" /tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc" /clone_lib="NIH_BMAP_FW0" /lab_host="DH10B" /note="Vector: pYX-ASC"
gene	1. .3844 /gene="Unc5a" /db_xref="LocusID:107448" /db_xref="MGI:894682"
CDS	252. .2780 /gene="Unc5a" /codon_start=1 /product="Unc5a protein" /protein_id="AAH58084.1"

```

/db_xref="GI:34784159"
/db_xref="LocusID:107448"
/db_xref="MGI:894682"
/translation="MAVRPGLWPALLGIVLTAWLRSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVCKAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEWQCQCVAWSSGTTKSQAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWRNEDLVDPSLDPNVYITREHSLVVRQARLADTPNYTCVAKNIVARRRSASAIVYVDGSWSPSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLLLLLVLVLYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPMLTIQPDLSTTTTYQGSILCPRDGSPKFQLSNHGHLSPLGSGRHTLHHSSPTSEAEDFVSRSLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCCEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPShLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLWKSLLVSYQEIPFYHIWNGTQQYLHCTFTLERVNASTSDLACKVWWQVEGDGQSFNFNITKDTRFAEMLALES EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPGNLGQLAAAVAGLGQPDAGLFTVSEAECC"
714. .926
/gene="Unc5a"
/note="IG; Region: Immunoglobulin"
/db_xref="CDD:smart00409"
984. .1130
/gene="Unc5a"
/note="TSP1; Region: Thrombospondin type 1 repeats"
/db_xref="CDD:smart00209"
1566. .1877
/gene="Unc5a"
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like netrin receptors"
/db_xref="CDD:smart00218"
2496. .2684
/gene="Unc5a"
/note="DEATH; Region: DEATH domain, found in proteins involved in cell death (apoptosis). Alpha-helical domain present in a variety of proteins with apoptotic functions. Some (but not all) of these domains form homotypic and heterotypic dimers"
/db_xref="CDD:smart00005"

```

ORIGIN

Query Match	72.6%	Score 1957.4;	DB 10;	Length 3844;
Best Local Similarity	84.6%	Pred. No. 0;		
Matches 2283;	Conservative 0;	Mismatches 246;	Indels 168;	Gaps 1;
Qy	1	ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGCCTCGCCGTTGGCTC	60	
Db	252	ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGCCTCACTGCCTGGCTT	311	
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120	
Db	312	CGTGGTTGGGTGCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCG	371	
Qy	121	GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA	180	
Db	372	GACCTGCTGCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAGCCA	431	

Qy	181	GTGCTGCTTGTGTCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCAACGGGAG 240
Db	432	GTGCTGCTGGTGTCAAGGCTGTGCCGCCACCCAGATCTTCTCAAGTGCAACGGGAA 491
Qy	241	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	492	TGGGTTGCCAGGTGATCACGTCAATTGAACGCAGCACTGACGGCAGCAGGGATTCCA 551
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db	552	ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTGAGAAAGTGTGGCTGGAG 611
Qy	361	GAATACTGGTGCCAGTGCAGTGGCATGGAGCTCCTGGGCCACCAAGAGTCAGAAGGCC 420
Db	612	GAGTACTGGTGCCAGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC 671
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db	672	TACATCCGGATTGCCTATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAAGTG 731
Qy	481	TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db	732	TCACTGGAGCAAGGCATTGTGCTACCTTGTGCCCCCGGAAGGAATCCCCCAGCTGAG 791
Qy	541	GTTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db	792	GTTGGAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATC 851
Qy	601	ACCGGGAGCACAGCTGGTGGCGACAGGCCGCCCTGCTGACACGGCCAATACACC 660
Db	852	ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCGCCCTGGCCGACACTCCAACTACACC 911
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGTCATCGTCTAC 720
Db	912	TGCGTGGCCAAGAACATCGTGGCCGTGCCGAAGGCCCTCTGGCCGCTATTGTTAT 971
Qy	721	GTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 780
Db	972	GTG----- 974
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCGGCCCTCTCAACGGGGCGCTTC 840
Db	975	----- 974
Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCCACCCGTGCCCAGTAGACGGCAGC 900
Db	975	----- GATGGGAGC 983
Qy	901	TGGAGCCCGTGGAGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGAGCCGT 960
Db	984	TGGAGCCCATGGAGTAAGTGGTCAGCCTGCCGGCTTGACTGCACCCACTGGCGAGCCGG 1043
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTGGAC 1020
Db	1044	GAGTGCTCCGACCCAGGCCGGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGAC 1103
Qy	1021	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGGCC 1080

Db	1104	ACCCGCAACTGTACCAGTGACCTCTGCCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCT	1163
Qy	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCCTGGTCCTGCTGCTGCTGGCTCATC	1140
Db	1164	CTCTACATCGGCCTCGTCGCCGTGGCCGTGCGCTCATCTGCTGCTGGTCCTCGTC	1223
Qy	1141	CTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1224	CTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGTGACTCATCCATCCTT	1283
Qy	1201	ACCTCAGGCTTCCAGCCCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1284	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTC	1343
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1344	ACCATCCAACCGGACCTCAGCACCAACGACCACCTACCAGGGCAGCTGTGTCCCCGG	1403
Qy	1321	CAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1404	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTAAATGGTCACCTGCTCAGCCCACGGC	1463
Qy	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTCTCC	1440
Db	1464	AGTGGCCGCCATACGCTGCACCACAGCTCCCCACCTCTGAGGCTGAGGACTTCGTCTCC	1523
Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1524	CGCCTCTCCACCCAAAACTACTTTCGTTCTGCCCGCGGTACAGAACATGGCTAT	1583
Qy	1501	GGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1584	GGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTC	1643
Qy	1561	ATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGACAAG	1620
Db	1644	ATACCCCCGGACGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAG	1703
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCATCGTTAGC	1680
Db	1704	CCAGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCTATCGTTAGC	1763
Qy	1681	TGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCTGGCTATGGACCAGTGTGGG	1740
Db	1764	TGTGGCCCCCAGGAGTCCTGCTACCCGGCCAGTCATCTGCCATGGACCAGTGCAGGG	1823
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1824	GAGCCCAGTCCCACAGCTGGAGCCTGCGCCTCAAAAGCAGTCCTGTGAGGGCAGCTGG	1883
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1884	GAGGACGTGCTGCACCTGGTGAGGAGTCGCCCTCATCTACTACTGCCAGCTGGAG	1943
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCTGGTGGAGAGGCC	1920

Db	1944	GCCGGGGCCTGCTATGTCTTACCGAGCAGCTAGGCCGTTGCCCTGGTGGAGAGGCC	2003
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCTCAAGCTGCTTGTTCAGGCTCCTGTACG	1980
Db	2004	CTCAGCGTGGCTGCCACCAAGGCCTCAGGCTCCTGTTCAGGCTGTACG	2063
Qy	1981	TCCCTCGAGTACAACATCCGGGCTACTGCCATGACACCCACGATGCACTCAAGGAG	2040
Db	2064	TCCCTCGAGTACAACATCCGAGTGTACTGCCATGACACCCACGATGCTCTCAAGGAG	2123
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCCTGCAC	2100
Db	2124	GTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCGTGTCCCTGCAC	2183
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2184	TTCAAAGACAGTTACCACAAACCTACGTCTGCCATCCACGACGTGCCAGCTCCCTGTGG	2243
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAACCTGCGCTACGCTACCAACGAGATCCCCTTTACACATCTGGAATGGCACG	2220
Db	2244	AAGAGCAAGCTCCTTGTCACTACCAACCTGCGCTACGCTACCAACGAGATCCCCTTTACACATCTGGAATGGACT	2303
Qy	2221	CAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2304	CAGCAGTATCTGCACCTTCACCCCTGGAGCGCGTCAATGCCAGCACCGACGACCTG	2363
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2364	GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTT	2423
Qy	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2424	AACATCACTAAGGACACGAGGTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGTC	2483
Qy	2401	CCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAACATAATT	2460
Db	2484	CCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCCTCCTCATTGGCAAAAGATCATT	2543
Qy	2461	TCCAGCCTGGACCCACCCCTGTAAGCGGGGTGCCACTGGCGGACTCTGGCCCAGAAACTC	2520
Db	2544	ACCAGCCTGGACCCACCCCTGCAGCCGGGGCGCGACTGGCGAACTCTAGCCCAGAAACTT	2603
Qy	2521	CACCTGGACAGCCATCTCAGCTCTTGCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db	2604	CACCTGGACAGCCATCTTAGCTCTTGCTCCAAGCCCAGCCCTACAGCCATGATCCTC	2663
Qy	2581	AACCTGTGGAGGCAGCGGGACTTCCCCAACGGAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Db	2664	AACCTATGGGAGGCAGCGGGACTTCCCCAACGGAACCTCGGCCAGCTGGCCAGCTGTG	2723
Qy	2641	GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	2724	GCCGGACTGGGCCAGCCAGATGCTGGCCTTCAACCGTGTCAAGAGGCCAGTGCTGA	2780

RESULT 11
BC009333

LOCUS BC009333 2688 bp mRNA linear PRI 25-MAR-2004
DEFINITION Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone IMAGE:4126760), partial cds.
ACCESSION BC009333
VERSION BC009333.2 GI:40226527
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2688)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2688)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On Dec 19, 2003 this sequence version replaced gi:14424611.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghghi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,

McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: g Column: 22.

FEATURES	Location/Qualifiers
source	1. .2688 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4126760" /tissue_type="Brain, neuroblastoma" /clone_lib="NIH_MGC_19" /lab_host="DH10B-R" /note="Vector: pOTB7"
gene	<1. .2688 /gene="UNC5A" /note="synonym: UNC5H1" /db_xref="LocusID:90249" /db_xref="MIM:607869"
CDS	<1. .1627 /gene="UNC5A" /codon_start=2 /product="UNC5A protein" /protein_id="AAH09333.2" /db_xref="GI:40226528" /db_xref="LocusID:90249" /db_xref="MIM:607869" /translation="DVALYVGLIAAVAVCLVLLLVILVYCRKKEGLDSVADSSILT SGFQPVSIKPSKADNPPLLTIQPDLSSTTTYQGSCLPRQDGSPSKFQLTNGHLLSPL GGGRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMPIPNTGI SLLIPPDAlPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILA MDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHYYCQLEASACYVFTEQLGR FALVGEALSVAAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ LIQEPRVLHFKDSYHNLRLSIHDVPSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT LERVSPSTSIDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS AFKIPFLIRQKIISSLDPGCRRGADWRTLQKLHLDHSFFASKPSPTAMILNLWEA RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"
misc_feature	413. .724 /gene="UNC5A" /note="ZU5; Region: Domain present in ZO-1 and Unc5-like netrin receptors" /db_xref="CDD:smart00218"
misc_feature	1343. .1594 /gene="UNC5A" /note="DEATH; Region: DEATH domain, found in proteins involved in cell death (apoptosis). Alpha-helical domain present in a variety of proteins with apoptotic functions. Some (but not all) of these domains form homotypic and heterotypic dimers" /db_xref="CDD:smart00005"

ORIGIN

Query Match 60.3%; Score 1625.4; DB 9; Length 2688;
Best Local Similarity 99.9%; Pred. No. 4.1e-269;

Qy	1911	GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTGCGCCGGT	1970
Db	841	GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTGCGCCGGT	900
Qy	1971	GCGCTGCACCTCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGATGC	2030
Db	901	GCGCTGCACCTCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGATGC	960
Qy	2031	ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG	2090
Db	961	ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG	1020
Qy	2091	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAG	2150
Db	1021	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAG	1080
Qy	2151	CTCCCTGTGGAAGAGTAAGCTCCTGTCAGCTACCAGGAGATCCCCTTTATCACATCTG	2210
Db	1081	CTCCCTGTGGAAGAGTAAGCTCCTGTCAGCTACCAGGAGATCCCCTTTATCACATCTG	1140
Qy	2211	GAATGGCACCGCAGCGGTACTTGCACTGCACCTTCACCCCTGGAGCGTGTAGCCCCAGCAC	2270
Db	1141	GAATGGCACCGCAGCGGTACTTGCACTGCACCTTCACCCCTGGAGCGTGTAGCCCCAGCAC	1200
Qy	2271	TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAG	2330
Db	1201	TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAG	1260
Qy	2331	CATCAACTCAACATACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGA	2390
Db	1261	CATCAACTCAACATACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGA	1320
Qy	2391	AGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATCGGCA	2450
Db	1321	AGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATCGGCA	1380
Qy	2451	GAAGATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCTGACTGGGGACTCTGGC	2510
Db	1381	GAAGATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCTGACTGGGGACTCTGGC	1440
Qy	2511	CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGC	2570
Db	1441	CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGC	1500
Qy	2571	CATGATCCTCAACCTGTGGAGGCAGCGGCACTTCCCCAACGGAACCTCAGCCAGCTGGC	2630
Db	1501	CATGATCCTCAACCTGTGGAGGCAGCGGCACTTCCCCAACGGAACCTCAGCCAGCTGGC	1560
Qy	2631	TGCAGCAGTGGCTGGACTGGCCAGCCAGACGCTGGCTCTTCACAGTGTGGAGGCTGA	2690
Db	1561	TGCAGCAGTGGCTGGACTGGCCAGCCAGACGCTGGCTCTTCACAGTGTGGAGGCTGA	1620
Qy	2691	GTCAGTGA 2697	
Db	1621	GTCAGTGA 1627	

RESULT 12

BD057525

LOCUS BD057525 1787 bp DNA linear PAT 27-AUG-2002
 DEFINITION Netrin receptors.
 ACCESSION BD057525
 VERSION BD057525.1 GI:22603131
 KEYWORDS JP 2001505062-A/2.
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE 1 (bases 1 to 1787)
 AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.,
 TITLE Netrin receptors
 JOURNAL Patent: JP 2001505062-A 2 17-APR-2001;
 THE REGENTS OF THE UNIV OF CALIFORNIA
 COMMENT PN JP 2001505062-A/2
 PD 17-APR-2001
 PF 19-FEB-1998 JP 1998536840
 PR 19-FEB-1997 US 08/808982
 PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
 PI MASU,
 PI KAZUKO KEINO MASU
 PC C07K1/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC
 Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES Location/Qualifiers
 source 1. .1787
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 57.6%; Score 1552.4; DB 6; Length 1787;
 Best Local Similarity 98.5%; Pred. No. 1.5e-256;
 Matches 1651; Conservative 0; Mismatches 16; Indels 9; Gaps 8;

Qy	1025	GCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT	1084
Db	1	GCAACTGTACCAAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT	59
Qy	1085	ATGTGGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTGCTTGTCCCTCATCCTCG	1144
Db	60	ATGTGGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTTGTCCCTCATCCTCG	119
Qy	1145	TTTATTGCCCGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT	1204
Db	120	TTTATTGCCCGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT	179
Qy	1205	CAGGCTTCAGCCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCT	1263
Db	180	CAGGCTTCAGCCCCGTCAGCATCTAACGCCAGCAAAGCAGACAACCCCCATCTGCTCACCT	239
Qy	1264	ATCCAGCCGGACCTCAGCACCAACCACCACTACCAAGGGCAGTCTCTGTCCCCGGCAG	1323
Db	240	ATCCAGCCGGACCTCAGCACCAACCACCACTACCAAGGGCAGTCTCTGTCCCCGGCAG	299

Qy	1324	GATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGTGGC	1383
Db	300	GATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGTGGC	359
Qy	1384	GGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCCCGC	1443
Db	360	GGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCCCGC	419
Qy	1444	CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGAACATGACCTATGGG	1503
Db	420	CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGAACATGACCTATGGG	479
Qy	1504	ACCTTCAACTCCTCGGGGGCGGCTGATGATCCCTAACAGGTATCAGCCTCCTCATC	1563
Db	480	ACCTTCAACTCCTCGGGGGCGGCTGATGATCCCTAACAGGAATCAGCCTCCTCATC	539
Qy	1564	CCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG	1623
Db	540	CCCCCAGATGCCATACCCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG	599
Qy	1624	GAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCCATCGTTAGCTGT	1683
Db	600	GAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCCATCGTTAGCTGT	659
Qy	1684	GGACCCCTGGCGTCTGCTCACCGGCCAGTCATCTGGCTATGGACCACTGTGGGAG	1743
Db	660	GGACCCCTGGCGTCTGCTCACCGGCCAGTCATCTGGCTATGGACCACTGTGGGAG	719
Qy	1744	CCCAGCCCTGACAGCTGGAGCCTGCGCTCAAAAGCAGTCGTGCGAGGGCAGCTGGGAG	1803
Db	720	CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAGCAGTCGTGCGAGGG-AGCTGGGAG	778
Qy	1804	GATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCC	1863
Db	779	GATGT-CTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGCC	837
Qy	1864	AGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGGCCCTGGTGGAGAGGCC	1923
Db	838	AGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGGCCCTGGTGGAGAGGCC	897
Qy	1924	AGCGTGGCTGCCGCCAAGCGCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACCTCC	1983
Db	898	AGCGTGGCTGCCGCCAAGCGCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACCTCC	957
Qy	1984	CTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAGGTG	2043
Db	958	CTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATGCACTCAAGGAGGTG	1017
Qy	2044	GTGCAGCTGGAGAACGAGCTGGGGGACAGCTGATCCAGGAGCACGGGTCTGCACCTC	2103
Db	1018	GTGCAGCTGGAGAACGAGCTGGGGGACAGCTGATCCAGGAGCACGGGTCTGCACCT-	1076
Qy	2104	AAGGACAGTTACCACAAACCTGCCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAG	2163
Db	1077	AAGGACAGTTACCACAAACCT--GCCCTATCCACGATGTGCCAGCTCCCTGTGGAAG	1134

RESULT 13

AX054976

LOCUS AX054976 9700 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 91 from Patent WO0073328.
ACCESSION AX054976
VERSION AX054976.1 GI:12228344
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS van Crielinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
TITLE Unc-5 constructs and screening methods
JOURNAL Patent: WO 0073328-A 91 07-DEC-2000;
Devgen NV (BE)
FEATURES Location/Qualifiers
source 1. .9700
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db xref="taxon:32630"

/note="plasmid pGC1037"

ORIGIN

Query Match 48.3%; Score 1302.8; DB 6; Length 9700;
Best Local Similarity 99.8%; Pred. No. 8.4e-214;
Matches 1304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1392 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCGCCTCTCAC 1451
Db 7 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCGCCTCTCAC 66

Qy 1452 CCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTATGGGACCTCAA 1511
Db 67 CCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTATGGGACCTCAA 126

Qy 1512 CTTCCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA 1571
Db 127 CTTCCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCAGA 186

Qy 1572 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAAGACGT 1631
Db 187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAAGACGT 246

Qy 1632 GAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1691
Db 247 GAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306

Qy 1692 TGGCGTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGGAGCCAGCCC 1751
Db 307 TGGCGTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGTGGGAGCCAGCCC 366

Qy 1752 TGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGGAGGATGTGCT 1811
Db 367 TGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGCTGGAGGATGTGCT 426

Qy 1812 GCACCTGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG 1871
Db 427 GCACCTGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG 486

Qy 1872 CTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCCCTCAGCGTGGC 1931
Db 487 CTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCCCTCAGCGTGGC 546

Qy 1932 TGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCGCCGTGGCCTGCACCTCCCTGAGTA 1991
Db 547 TGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCGCCGTGGCCTGCACCTCCCTGAGTA 606

Qy 1992 CAACATCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT 2051
Db 607 CAACATCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT 666

Qy 2052 GGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCTGCACTCAAGGACAG 2111
Db 667 GGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCTGCACTCAAGGACAG 726

Qy 2112 TTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT 2171
Db 726 TTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTAAGCT 2171

Db	727 TTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGAAGAGTAAGCT	786
Qy	2172 CCTTGTCAAGCTACCAGGAGATCCCCTTTATCACATCTGAATGGCACGCAGCGGTACTT	2231
Db	787 CCTTGTCAAGCTACCAGGAGATCCCCTTTATCACATCTGAATGGCACGCAGCGGTACTT	846
Qy	2232 GCACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCTGCAAGCT	2291
Db	847 GCACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTGGCTGCAAGCT	906
Qy	2292 GTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACATCACCAA	2351
Db	907 GTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACATCACCAA	966
Qy	2352 GGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAAGCBBBBBCCAGCCCTGGT	2411
Db	967 GGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAAGCBBBBBCCAGCCCTGGT	1026
Qy	2412 GGGCCCCAGTGCCTCAAGATCCCCTCATTGGCAGAAGATAATTCCAGCCTGGA	2471
Db	1027 GGGCCCCAGTGCCTCAAGATCCCCTCATTGGCAGAAGATAATTCCAGCCTGGA	1086
Qy	2472 CCCACCCCTGTAGGCBBBBBCCAGTGGGGACTCTGGCCAGAAAACCTCCACCTGGACAG	2531
Db	1087 CCCACCCCTGTAGGCBBBBBCCAGTGGGGACTCTGGCCAGAAAACCTCCACCTGGACAG	1146
Qy	2532 CCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAGCCATGATCCTAACCTGTGGGA	2591
Db	1147 CCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAGCCATGATCCTAACCTGTGGGA	1206
Qy	2592 GGCGCGGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGG	2651
Db	1207 GGCGCGGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGG	1266
Qy	2652 CCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	1267 CCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGCTGA	1312

RESULT 14

MMU72634

LOCUS	MMU72634	9299 bp	mRNA	linear	ROD	13-MAY-1997
DEFINITION	Mus musculus rostral cerebellar malformation protein (rcm) mRNA, complete cds.					
ACCESSION	U72634					
VERSION	U72634.1 GI:2088526					
KEYWORDS	.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE	1 (bases 1 to 9299)					
AUTHORS	Ackerman, S.L., Kozak, L.P., Przyborski, S.A., Rund, L.A., Boyer, B.B. and Knowles, B.B.					
TITLE	The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein					
JOURNAL	Nature 386 (6627), 838-842 (1997)					

MEDLINE 97271898
 PUBMED 9126743
 REFERENCE 2 (bases 1 to 9299)
 AUTHORS Ackerman, S.L., Kozak, L.P., Rund, L.A. and Knowles, B.B.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
 Bar Harbor, ME 04609, USA
 FEATURES Location/Qualifiers
 source 1. .9299
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57B6/SJL"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="between Amyl and Mpmv9"
 gene 1. .9299
 /gene="rcm".
 CDS 151. .2946
 /gene="rcm"
 /function="transmembrane receptor"
 /note="similar to C. elegans UNC-5"
 /codon_start=1
 /product="rostral cerebellar malformation protein"
 /protein_id="AAB54103.1"
 /db_xref="GI:2088527"
 /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE
 FFHELPETFPSPDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
 KDHVVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSAGTTKSRAKAY
 RIAYLRKTFEQEPLGKEVSLEQEVLQLQCRPEGIPVAEVEWLKNEDIIIDPAEDRNFYI
 TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
 GRGYQKRRTCTNPAPLNGGAFCEGQSVKIACTLCPVDGRWTSWSKWSTCGTECH
 WRRRECTAPAKNNGGKDCDGLVLQSKNCTDGLCMQAAPDSDVALYVGIVIAVTVCLA
 ITVVVALFVYRKNRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAMYR
 GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
 LLENEALNLKNQSLARQTDPSCTAFGTFSNLGGHLIIPNSGVSLIIPAGAIPOQGRVYE
 MYVTVHRKENMRPMPEDSQTLTVPVSCGPGALLTRPVILTLHHCADPSTEDWKIQL
 KNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHILLENSTYALVGQSTTCAAALKL
 KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
 LRLSIHDIAHSLWKSKLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
 VRQVEGEQGQFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPPLIRQKLCSSL
 DAPQTRGHDRMLAHKLNDRYLNKFATKSSPTGVILDLWEAQNFPDGNLMSLAAVLE
 EMGRHETVVS LAAEGQY"

ORIGIN

Query Match 36.8%; Score 992; DB 10; Length 9299;
Best Local Similarity 62.7%; Pred. No. 1.9e-160;
Matches 1660; Conservative 0; Mismatches 925; Indels 63; Gaps 5;

Db	428	TCAAGTCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG	487
Qy	284	GGAGCAGTGGCTGCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGA	343
Db	488	AAACCTCTGGTCTAATTGTGAGAGAAAGTGAGCATGAGATTCAAGCCAGCAGGTGGAGG	547
Qy	344	AGGTGTTGGGCTGGAGGAATACTGGTGCAGTGCAGTCGGATGGAGCTCCTGGCACCA	403
Db	548	AACTGTTGGGCTGAAGATTACTGGTGCAGTGTGGCCTGGAGCTCAGCAGGCACTA	607
Qy	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCAGAACACTCGAGCAGGAGC	463
Db	608	CGAAGAGTCAGGACATACGTGCATTGCGTATCTCGGAAGACATTGAGCAGGAAC	667
Qy	464	CGCTGGCCAAGGAGGTGTCCTGGAGCAGGGCATCGTGCCTGCCGCCACCGGAGG	523
Db	668	CCTTGGGAAAGGAAGTGTCCCTGGAGCAGGAAGTCTTACTCCAGTGTGGCACCTGAAG	727
Qy	524	GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGACCCGTCCTGG	583
Db	728	GGATCCCAGTGGCTGAGGTGGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG	787
Qy	584	ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCTTGTG	643
Db	788	ATCGGAACCTTATATTACTATCGATCACAAACCTGATCATCAAGCAAGCCCCACTCTCAG	847
Qy	644	ACACGGCCAACCTACACCTCGTGGCCAAGAACATCGTGGCACGTGCCGCAGCGCCTCCG	703
Db	848	ATACAGCAAATTATACCTGTGTTGCCAAAATTGTTGCCAGAGAAAAAGCACACAG	907
Qy	704	CTGCTGTCATCGTCTACGTGAACGGTGGTGGTCACGTGGACCGAGTGGCCGTCTGCA	763
Db	908	CCACTGTCATCGTGTATGTTAATGGTGGCTGGCACCTGGACAGAGTGGTCTGTGTGTA	967
Qy	764	GCGCCAGCTGTGGCGCGCTGGCAGAACCGAGCCGGAGCTGCACCAACCCGGCGCTC	823
Db	968	ACAGCCGCTGTGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC	1027
Qy	824	TCAACGGGGCGCTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCCACCTGT	883
Db	1028	TCAATGGTGGGCCTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT	1087
Qy	884	GCCCAGTAGACGGCAGCTGGAGCCGTGGAGCAAGTGGTCGGCTGTGGACTGCA	943
Db	1088	GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGACTGAATGCA	1147
Qy	944	CCCACGGCGAGCCGTGAGTGCCTGCACCCAGCACCCCGCAACGGAGGGAGGAGTGC	1003
Db	1148	CCCACGGCGAGGAGGGAGTGTACAGCACCAAGAACGGGGTAAGGACTGTG	1207
Qy	1004	AGGGCACTGACCTGGACACCCGAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTG	1063
Db	1208	ATGGCCTGGCCTCCAATCCAAGAACTGCACTGATGGCTGTGCATGCAGGCTGCTCCTG	1267
Qy	1064	GCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATGCCGTGGCGTCTGCCTGGTCC	1120

Db	1268	ACTCAGATGATGGCTCTACGTGGGATTGTGATCGCTGTAACAGTCTGTCTGGCGA	1327
Qy	1121	TGCTGCTGCTTGCCTCATCCTCGTTATTGCCGAAAGAAGGAGGGCTGGACTCAGATG	1180
Db	1328	TCACTGTTGTGGTGGCCCTGTTGTATCGGAAGAACACCCTGACTTGAGTCTGACA	1387
Qy	1181	TGGCTGACTCGTCATTCTCACCTCAGGCTTCCAGCCCCTCAGCATCAAGCCCAGCAAAG	1240
Db	1388	TCATTGACTCCTCAGCACTCAATGGCGCTTTCAGCCTGTGAACATCAAG-----G	1438
Qy	1241	CAGACAAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCTACC	1300
Db	1439	CTGCCAGACAAGATCTCCTGGCTGTCCCCCTGACCTCACCTCAGCTGCAGCCATGTACA	1498
Qy	1301	AGGGCAGTCTCTGCCCCGGCAGGATG-----GGCCCAGCCCAAGTCCAGCTCACCAA	1355
Db	1499	GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCAATGACCAACTCTCAA	1558
Qy	1356	TGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCACACACTGCACCAAGCTCTCCCAC	1415
Db	1559	TTCTGGACCCACTACCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCAC	1618
Qy	1416	CTC---TGAGGCCGAGGAGTTCGTCTCCGCCCTCCACCC-----	1453
Db	1619	CTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTACCCCCAGATGACCCAGTCCTTGC	1678
Qy	1454	-----AGAACTACTTCCGCTCCCTGCCCGAGGCACCAAGCA	1489
Db	1679	TAGAGAATGAGGCCCTAACCTGAAGAACCAAGAGCCTCGCAAGACAGACTGACCCATCCT	1738
Qy	1490	ACATGACCTATGGACCTTCAACTTCCCTGGGGGGCGGCTGATGATCCCTAACAGGTA	1549
Db	1739	GCACAGCATTGGTACCTCAACTCTTGGGGTCACCTCATCATTCTAACAGGAG	1798
Qy	1550	TCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCA	1609
Db	1799	TAAGCTTGCTGATTCCCGCTGGGCCATTCTCAGGGGAGAGTCTATGAAATGTATGTGA	1858
Qy	1610	CGCTGCACAAGCCGGAAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTC	1669
Db	1859	CTGTACACAGGAAGAAAATATGAGGCCCTCATGGAAGACTCTCAGACCCCTACTAACCC	1918
Qy	1670	CCATCGTTAGCTGTGGACCCCTGGCGCTCTGCTCACCCGCCAGTCATCCTGGCTATGG	1729
Db	1919	CTGTGGTGAGCTGTGGCCTCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGC	1978
Qy	1730	ACCACTGTGGGAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCG	1789
Db	1979	ATCACTGTGCAGACCCAGCACCGAGGAACAGTCCAGCTCAAAACCAGGCAGTGC	2038
Qy	1790	AGGGCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACT	1849
Db	2039	AGGGACAATGGGAGGATGTTGTGGTGGGAGGAACCTCACACCCCCCTGTTACA	2098
Qy	1850	GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCTGG	1909
Db	2099	TTCAGCTGGATGCAGAGGCTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGG	2158

Qy	1910	TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGTTCTGTTGCCGG 1969
Db	2159	TTGGGCAGTCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTGCCCC 2218
Qy	1970	TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCATGACACCCACGATG 2029
Db	2219	TCTGCTGCTCTCCCTGGAGTACAGCATTAGAGTCTACTGCCCTGGATGACACACAGGATG 2278
Qy	2030	CACTCAAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCAC 2089
Db	2279	CCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGAGGACAGCTCCTAGAAGAACCA 2338
Qy	2090	GGGTCCCTGCACCAAGGACAGTTACCACAAACCTGCCCTATCCATCCACGATGTGCCCA 2149
Db	2339	AGGCTCTTCATTTAAAGGCAGCATCCACAACCTGCCCTGTCTATTGACATCGCCC 2398
Qy	2150	GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAAGCTTACCAACCTGCCCTATCCATCCACGATGTGCCCA 2209
Db	2399	ATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTACACATCT 2458
Qy	2210	GGAATGGCACGCAGCGGTACTTGCACCTCACCCCTGGAGCGTGTCAAGCCCCAGCA 2269
Db	2459	GGAGTGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGAAAGACTCAGCCTAAACA 2518
Qy	2270	CTAGTGACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA 2329
Db	2519	CAGTGGAACTGGTTGCAAACCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTCC 2578
Qy	2330	GCATCAACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTG 2389
Db	2579	AGCTCAACTGTACTGTGTCAAGAGAACCTACTGGCATCGACTTACCTCTCCCTGGACCCTG 2638
Qy	2390	AAGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATCGC 2449
Db	2639	CTAGTACCATCACCACTGTACCGGACCAAGTGCTTCAGCATTCCCTCCCTATCCGGC 2698
Qy	2450	AGAAGATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGG 2509
Db	2699	AGAAGCTATGCAGCAGCCTGGATGCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGG 2758
Qy	2510	CCCAGAAAATCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAG 2569
Db	2759	CCCATAAAACTCAACCTGGACAGGTACTTGAATTACTTGCCACCAAATCGAGCCCAACTG 2818
Qy	2570	CCATGATCCTCAACCTGTGGAGGCAGCGCCACTTCCCCAACGGCAACCTCAGCCAGCTGG 2629
Db	2819	GCGTAATCCTGGATCTTGGAAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGG 2878
Qy	2630	CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTG 2689
Db	2879	CAGCCGTCCCTGGAAGAAATGGAAGACATGAGACAGTGGTGTCCCTGGCAGCAGAAGGAC 2938
Qy	2690	AGTGCTGA 2697
Db	2939	
Qy	2946	AGTATTGA 2946

RESULT 15

AY187310

LOCUS AY187310 2962 bp mRNA linear VRT 06-JUN-2003
 DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.
 ACCESSION AY187310
 VERSION AY187310.1 GI:31442350
 KEYWORDS .
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 2962)
 AUTHORS Guan,W. and Condic,M.L.
 TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression
 during chick dorsal root ganglia development
 JOURNAL Gene Expr. Patterns 3, 369-373 (2003)
 REFERENCE 2 (bases 1 to 2962)
 AUTHORS Guan,W. and Condic,M.L.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah,
 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA
 FEATURES Location/Qualifiers
 source 1. .2962
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 CDS 1. .2796
 /codon_start=1
 /product="UNC5-like protein 3"
 /protein_id="AA067275.1"
 /db_xref="GI:31442351"
 /translation="MGKGLEGTAARCGLMGYLLHSVLPALAVLGASRPGSAAQDDD
 FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
 KDHVVDERVDETSGLIVCEVSIEISRQQVEELFGPEDYWCQCVAWSAGTTKSRKAYV
 RIAYLRKTFEQEPLGKEVSLEQEVLQLCRRPEGIPVAEVEWLKNEEVIDPVEDRNFYI
 TIDHNLIQKARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC
 GRGFQKRTRTCTNPAPLNGGAFCEGQNVQKIACTTLCVPDGKWTWSWKSTCGTECTH
 WRRRECTAPAKPNNGGDCEGLVLSKNCTDGLCMQAAPDSDDVALYVGIVIAVIVCLA
 ISVVVALFVYRKNH RDFESDIIDSSALNGFQPVNKAARQDLLAVPPDLTSAAMYR
 GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS
 LLENETLNVKNQSLARQTDPSCATAFGTFNSLGGHLVIPNSGVSLIIPAGAVPQGRVYE
 MYVTVHRKEGMRPPVEDSQTLT PVVSCGGPGALLTRPVVLTMHCAEPNMDDWQIQQL
 KHQAGQGPWEDVVVGEENFTTPCYIQLDPEACHILTELSTYALVGQSITKAAKRL
 KLAIFGPLCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLEEPKTLHFKGSTHN
 LRLSIHDIAHSLWKS KLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSINTLELVCKLC
 VRQVEGEQGI FQLNCVSEEPTGIDYPIMDSAGSITIVGPNAFSIPLPIRQKLCSSL
 DAPQTRGHDRMLAHKLKDRLNYFATKSSPTGVILDWEAQNF PDGNLSMLAAVLE
 EMGRHETVVSLAAEGNY"

ORIGIN

Query Match 36.7%; Score 988.8; DB 5; Length 2962;
 Best Local Similarity 62.6%; Pred. No. 8e-160;
 Matches 1658; Conservative 0; Mismatches 927; Indels 63; Gaps 5;

QY 104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163

Db	158	TTCCCTTCTGATCCTCCAGAGCCATTGCCCACTTCTCATTGAACCGAAGAAGCTTACA	217
Qy	164	TCGTCAAGAACAAAGCCAGTGCCTGCTGTGCAAGGCCGTGCCGCCACGCAGATCTCT	223
Db	218	TCGTAAAAAACAGCCTGTGAATCTGTACTGCAAAGCGAGCCCTGCCACGCAGATCTATT	277
Qy	224	TCAAGTGCACGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCCAGCACAGACG	283
Db	278	TTAAGTCAACAGTGAATGGGTTCATCAGAAGGATCATGTGGTGGATGAGAGAGTAGATG	337
Qy	284	GGAGCAGTGGCTGCCACCAGGAGTCCGATTAATGTCAGGCAAGCAGCAGGTCGAGA	343
Db	338	AAACCTCTGGTCTGATCGTCTGTGAGGTGAGCATCGAGATTCCCAGCAGGTGGAAG	397
Qy	344	AGGTGTTCGGGCTGGAGGAATACTGGTGCAGTGCCTGGCATGGAGCTCCCGACCCA	403
Db	398	AGCTCTTGAGCCGAGGACTACTGGTGCAGTGTGTCGCTGGAGCTCAGCTGGCACCA	457
Qy	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC	463
Db	458	CCAAGAGCCGCAAGGCCTACGTCCGCATTGCATATCTCAGAAAGACTTTGAGCAGGAGC	517
Qy	464	CGCTGGCCAAGGAGGTGTCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGG	523
Db	518	CGCTGGGAAAGAAGTGTCCCTGGAGCAAGAGGTCTGCTCAGTGCCTCCCTGAAG	577
Qy	524	GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCCTGG	583
Db	578	GCATTCCAGTAGCTGAGGTAGAGTGGCTGAAGAAATGAAGAGGTGATCGATCCTGTGGAAG	637
Qy	584	ACCCCCAATGTATAACATCACCGGGAGCACAGCCTGGTGGCGACAGGCCGCCCTGCTG	643
Db	638	ACCGAAATTTCACATCACCATTGATCACAAACCTGATCATCAAGCAAGCCGGCTTCCG	697
Qy	644	ACACGGCCAACCTACACCTGCGTGGCAAGAACATCGTGGCACGTCGCCAGCGCCTCCG	703
Db	698	ACACGGCTAACTACACCTGTGTGCCAAAACATTGTGGCAAAAGGAAAAGCACGACAG	757
Qy	704	CTGCTGTCATCGTCTACGTGAACGGTGGTGGCGACGTGGACCGAGTGGTCCGCTGCA	763
Db	758	CAACTGTGATTGTCTATGTGAATGGAGGCTGGCTACCTGGACCGAGTGGTACCGTGCA	817
Qy	764	GCGCCAGCTGTGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCGGCCCTC	823
Db	818	ACAGCCGCTGTGGAGAGGCTCCAGAAGCGCACAAGGACCTGCACTAACCTGCCAAC	877
Qy	824	TCAACGGGGCGCTTCTGTGAGGGCAGAACATGTCAGAAAACAGCCTGCCACCCCTGT	883
Db	878	TCAATGGGGGGCCTCTGCAGAGGGCAAAATGTTAGCTGCACCAACCCCTGT	937
Qy	884	GCCCAGTAGACGGCAGCTGGAGCCGTGGAGCAAGTGGTGGCCTGTGGACTGCA	943
Db	938	GTCCAGTGGATGGCAAATGGACGTCCCTGGAGCAAGTGGCACTTGTGGCACAGAGTGTA	997
Qy	944	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCC	1003

Db	998	CCCACTGGCGCCGGAGGGAGTGCACAGCTCCGGCCCCGAAGAATGGAGGCAAGGACTGTG	1057
Qy	1004	AGGGCACTGACCTGGACACCCGCAACTGTACCATGACCTCTGTACACAGTGCTCTG	1063
Db	1058	AGGGACTGGTGCAGTCTAAAGAACTGCACGTGATGGCTCTGCATGCAGGCTGCACCTG	1117
Qy	1064	GCCCTGAGGACGTGGCCCTATGTGGG---CCTCATGCCGTGGCGTGCCTGGTCC	1120
Db	1118	ACTCGGATGATGTTGCTCTACGTGGGGATTGTCTATTGCTGTGATTGTGTGCCTGGCTA	1177
Qy	1121	TGCTGCTGCTTGTCCATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATG	1180
Db	1178	TTTCTGTGGTTGTGGCCCTGTTGTCTATCGCAAGAACCAACCGTGACTTGAGTCAGATA	1237
Qy	1181	TGGCTGACTCGTCATTCTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAG	1240
Db	1238	TTATCGACTCATCGCGCTAAATGGGGATTTCAGCCTGTTAACATCAAGGCTGCAAGAC	1297
Qy	1241	CAGACAACCCCCATCTGCTACCCTCAGGCCGGACCTCAGCACCAACCAACCTACC	1300
Db	1298	AAGAC-----CTCTTGGCAGTGCCACCAGACCTCACTCTGCTGCAGCCATGTACA	1348
Qy	1301	AGGGCAGTCTCTGTCGGCAGGATG----GCCAGCCCCAAGTCCAGCTCACCAA	1355
Db	1349	GGGGGCCTGTGTATGCCCTGCATGATGTCTCTGATAAAATCCAATGACCAATTCTCCGA	1408
Qy	1356	T----GGCACCTGCTCAGCCCCCTGGTGGCGGGCCACACACTGCACCAAGCTCTC	1411
Db	1409	TCCTGGACCCACTGCCAATCTGAAGATTAAAGTTATAACACCTCTGGAGCAGTCACCC	1468
Qy	1412	CCACCTCTGAGGCCGAGGAGTTCGTCTCCGCCCTCCACCC-----	1453
Db	1469	CCCAGGATGAACCTCTGACTTCTCCCTCAAGCTGTCCCCACAGATTACCCAGTCTGT	1528
Qy	1454	-----AGAACTACTTCCGCTCCCTGCCCGAGGCACCAAGCA	1489
Db	1529	TGGAGAAATGAGACTCTGAACGTGAAGAACCAAGCCTTGACGGCAAACAGACCCATCCT	1588
Qy	1490	ACATGACCTATGGGACCTTCAACTTCTCGGGGGCCGGCTGATGATCCCTAACAGGTA	1549
Db	1589	GCACTGCATTGGGACCTTCAACTCGTTAGGGGGCACCTAGTAATTCTAACCTGAGGAG	1648
Qy	1550	TCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCA	1609
Db	1649	TGAGCTGCTGATCCCAGCAGGGGCTGTTCCCCAAGGAAGAGTCTATGAAATGTATGTGA	1708
Qy	1610	CGCTGCACAAGCCGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCCTGCTGAGTC	1669
Db	1709	CAGTCCACAGGAAGGGAGGGCATGAGACCAACCTGTAGAAGACAGCCAGACGCTGACAC	1768
Qy	1670	CCATCGTTAGCTGGACCCCTGGCGTCCTGCTCACCCGCCAGTCATCCTGGCTATGG	1729
Db	1769	CAGTGGTGAGCTGTGGCCCACCAGGAGCGCTGCTGACCCGACCCGTTGTGCTGACCATGC	1828
Qy	1730	ACCACTGTGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCG	1789
Db	1829	ACCACTGTGCTGAGCCCCAACATGGATGACTGGCAGATCCAGCTCAAGCACCAGGAGGCC	1888

Qy	1790	AGGGCAGCTGGGAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTACTACT	1849
Db	1889	AGGGACCAGTGGGAGGATGTAGTGGTGGCGGGAGGAAACTTCACCACTCCATGCTACA	1948
Qy	1850	GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCCTGG	1909
Db	1949	TCCAGCTGGACCCAGAGGCCTGTCATATCCTGACGGAGACCCCTCAGCACGTACGCCCTGG	2008
Qy	1910	TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGGCCCTCAAGCTGCTTGTGCGCCGG	1969
Db	2009	TGGGACAATCCATACCAAAGCAGCAGGCCAACGCTCTCAAATTGCCATCTTGAGGACAC	2068
Qy	1970	TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCATGACACCCACGATG	2029
Db	2069	TGTCCTGTTCCCTCACTGGAGTACAGCATCCCGCTACTGCCATGACACACAGGATG	2128
Qy	2030	CACTCAAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGGCCAC	2089
Db	2129	CCCTGAAGGAGGTCCCTCCAGCTTGAGCGGAGATGGTGGCAGCTGTTGGAGGAACCCA	2188
Qy	2090	GGGTCCCTGCACTTAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCCA	2149
Db	2189	AAACTTGCATTTAAGGAAAGTACCCACAACCTGCGCTATCCATTGACATTGCC	2248
Qy	2150	GCTCCCTGTGAAAGAGTAAGCTCCTGTCAGCTACCAGGAGATCCCCTTTATCACATCT	2209
Db	2249	ACTCTCTCTGGAAGAGCAAACGCTGCCGCTAAATACCAGGAGATTCCCTTACACATCT	2308
Qy	2210	GGAATGGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCA	2269
Db	2309	GGAGTGGTGCCAGAGGAACCTGCACTGCACCTCACGCTGGAACGATTCACTCAATA	2368
Qy	2270	CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTCA	2329
Db	2369	CCCTGGAGCTCGTCTGCAAACCTGTGTCGGCAAGTCGAAGGAGAAGGGCAGATCTTCC	2428
Qy	2330	GCATCAACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTG	2389
Db	2429	AGCTGAAGTGCCTCAGTATCAGAGGAACCCACTGGCATTGATTATCCCACATGGATTCA	2488
Qy	2390	AAGCGGGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGC	2449
Db	2489	CAGGCAGCATCACAAACGATAGTTGGGCCAACGCTTCAGCATCCCCCTCCAATAAGGC	2548
Qy	2450	AGAAGATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGG	2509
Db	2549	AGAAGCTCTGCAGCAGCCTGGATGCACCCAGACCCGGGCCATGACTGGAGGATGCTGG	2608
Qy	2510	CCCAGAAAATCCACCTGGACAGCCATCTCAGCTTGTGCTCCAAGGCCAGCCCCACAG	2569
Db	2609	CCCACAAGCTGAAATTGGACAGGTACCTAAATTATTTGCTACGAAGTCGAGTCCACTG	2668
Qy	2570	CCATGATCCTCAACCTGTGGAGGCGGGCAGTCCCAACGGCAACCTCAGCCAGCTGG	2629
Db	2669	GGGTGATCCTGGATCTGGAGGCCAGAATTCCCTGATGGCAACCTGAGCATGCTGG	2728

Search completed: March 6, 2005, 05:25:12
Job time : 11476.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 14:42:51 ; Search time 1373 Seconds
(without alignments)
11628.216 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002as:
7: geneseqn2002bs:
8: geneseqn2003as:
9: geneseqn2003bs:
10: geneseqn2003cs:
11: geneseqn2003ds:
12: geneseqn2004as:
13: geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
1	2697	100.0	2752	6	ABK37922	Abk37922 cDNA enco
2	2697	100.0	2752	12	ADH71617	Adh71617 Human gen
3	2687.4	99.6	2697	6	ABK52891	Abk52891 Human net
4	2621.4	97.2	2881	6	ABK49422	Abk49422 DNA encod

5	2621.4	97.2	2881	12	ADH71609	Adh71609 Human gen
6	2619.8	97.1	2880	12	ADH71633	Adh71633 Human gen
7	2619.8	97.1	2881	12	ADH71649	Adh71649 Human gen
8	2619.8	97.1	2881	12	ADH71635	Adh71635 Human gen
9	2619.8	97.1	2881	12	ADH71637	Adh71637 Human gen
10	2619.8	97.1	2881	12	ADH71641	Adh71641 Human gen
11	2619.8	97.1	2881	12	ADH71629	Adh71629 Human gen
12	2619.8	97.1	2881	12	ADH71631	Adh71631 Human gen
13	2619.8	97.1	2881	12	ADH71645	Adh71645 Human gen
14	2619.8	97.1	2881	12	ADH71627	Adh71627 Human gen
15	2619.8	97.1	2881	12	ADH71639	Adh71639 Human gen
16	2619.8	97.1	2881	12	ADH71643	Adh71643 Human gen
17	2619.8	97.1	2881	12	ADH71625	Adh71625 Human gen
18	2619.8	97.1	2881	12	ADH71647	Adh71647 Human gen
19	2618.2	97.1	2881	10	ADG42568	Adg42568 Novel hum
20	2368.6	87.8	2907	4	AAK52261	Aak52261 Human pol
21	2343	86.9	3561	12	ADL06497	Adl06497 Human tum
22	2343	86.9	3580	6	ABK15169	Abk15169 Human REP
23	2318	85.9	2575	12	ADH71621	Adh71621 Human gen
24	2266.2	84.0	2463	12	ADH71623	Adh71623 Human gen
25	2252.2	83.5	2697	6	AAS16843	Aas16843 Rat netri
26	2252.2	83.5	3014	2	AAV52940	Aav52940 Rat UNC-5
27	1625.4	60.3	2635	11	ADN95100	Adn95100 Human LEC
28	1552.4	57.6	1787	2	AAV52941	Aav52941 Human UNC
29	1302.8	48.3	9700	4	AAC90958	Aac90958 Plasmid p
30	1200.6	44.5	1321	4	AAH99530	Aah99530 Human pro
31	987.6	36.6	1002	12	ADH71619	Adh71619 Human gen
32	985.6	36.5	1009	12	ADH71613	Adh71613 Human gen
33	970	36.0	3646	5	AAS75738	Aas75738 DNA encod
34	970	36.0	3646	13	ADR99257	Adr99257 Human unc
35	968.8	35.9	2796	10	AAL56266	Aal56266 Human thr
36	943.4	35.0	993	12	ADH71611	Adh71611 Human gen
37	936.2	34.7	2860	6	ABT06279	Abt06279 Human NOV
38	936.2	34.7	2860	6	ABT06280	Abt06280 Human NOV
39	912.2	33.8	2895	6	ABQ93898	Abq93898 Human tra
40	904.2	33.5	2986	13	ADR99249	Adr99249 Human 1RO
41	904.2	33.5	3933	13	ADR07892	Adr07892 Full leng
42	901	33.4	3884	4	AAS21316	Aas21316 Human cDN
43	901	33.4	3884	8	ACA03675	Aca03675 cDNA enco
44	901	33.4	3884	8	ABX89213	Abx89213 DNA encod
45	901	33.4	3884	8	ACD41867	Acd41867 Human sec

ALIGNMENTS

RESULT 1

ABK37922

ID ABK37922 standard; cDNA; 2752 BP.

XX

AC ABK37922;

XX

DT 21-MAY-2002 (first entry)

XX

DE cDNA encoding Human protein NOV1.

XX

KW Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;

KW cell signal processing disorder; metabolic disorder; obesity; infection;
KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW psychosis; neurological disorder; anxiety; schizophrenia;
KW manic depression; dementia; dyskinesia; Huntington's disease;
KW Gilles de la Tourette's syndrome; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200210216-A2.

XX

PD 07-FEB-2002.

XX

PF 30-JUL-2001; 2001WO-US024225.

XX

PR 28-JUL-2000; 2000US-0221409P.
PR 04-AUG-2000; 2000US-0222840P.
PR 04-AUG-2000; 2000US-0223752P.
PR 04-AUG-2000; 2000US-0223762P.
PR 04-AUG-2000; 2000US-0223769P.
PR 04-AUG-2000; 2000US-0223770P.
PR 14-AUG-2000; 2000US-0225146P.
PR 15-AUG-2000; 2000US-0225392P.
PR 15-AUG-2000; 2000US-0225470P.
PR 16-AUG-2000; 2000US-0225697P.
PR 01-FEB-2001; 2001US-0263662P.
PR 05-APR-2001; 2001US-0281645P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;

XX

DR WPI; 2002-180074/23.
DR P-PSDB; AAU85403.

XX

PT New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT immune disorders.

XX

PS Claim 9; Page 9-10; 213pp; English.

XX

CC The invention relates to an isolated cytoplasmic, nuclear, membrane
CC bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC form. Also included are the nucleic acids encoding the NOVX proteins, a
CC vector comprising the nucleic acid, a cell comprising the vector, an anti
CC -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC antibody are useful for treating or preventing a NOVX-associated
CC disorder, where the disorder is selected from cardiomyopathy,
CC atherosclerosis, diabetes, a disorder related to cell signal processing
CC and metabolic pathway modulation, metabolic disorders, obesity,
CC infectious disease, anorexia, cancer-associated cachexia, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,

CC immune disorders, haematopoietic disorders, and the various
CC dyslipidaemias, metabolic disturbances associated with obesity, the
CC metabolic syndrome X and wasting disorders associated with chronic
CC diseases, bacterial, fungal, protozoal and viral infections, pain,
CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC hypertrophy, and psychotic and neurological disorders, including anxiety,
CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,
CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC nucleic acid is useful in gene therapy. The present sequence encodes a
CC NOVX protein

XX

SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2697; DB 6; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGTTGGCTC 60	
Db	46 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGCATAGTCCTCGCCGTTGGCTC 105	
Qy	61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120	
Db	106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165	
Qy	121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180	
Db	166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 225	
Qy	181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 240	
Db	226 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 285	
Qy	241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300	
Db	286 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 345	
Qy	301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360	
Db	346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 405	
Qy	361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCCACCAAGAGTCAGAAGGCC 420	
Db	406 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCCACCAAGAGTCAGAAGGCC 465	
Qy	421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 480	
Db	466 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGGAGTG 525	
Qy	481 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540	
Db	526 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585	
Qy	541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAAATGTATACATC 600	

Db	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGCCCTGGACCCAATGTATACATC	645
Qy	601	ACGCGGGAGCACAGCTGGTGGCGACAGGCCGCCCTGCTGACACGGCCAACATACACC	660
Db	646	ACGCGGGAGCACAGCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCCAACATACACC	705
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	720
Db	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC	765
Qy	721	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	766	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	825
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	840
Db	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	885
Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	900
Db	886	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	945
Qy	901	TGGAGCCCCTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	946	TGGAGCCCCTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	1006	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1125
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGCCTGCTGCTGCTGTGCTCATC	1140
Db	1126	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGCCTGCTGCTGCTGTGCTCATC	1185
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1186	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1306	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGT	1380
Db	1366	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGT	1425
Qy	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCC	1440
Db	1426	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTCTCC	1485

Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1486	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1545
Qy	1501	GGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCC	1560
Db	1546	GGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCC	1605
Qy	1561	ATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1606	ATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Db	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1725
Qy	1681	TGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1726	TGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1785
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1846	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1905
Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1920
Db	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1965
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC	1980
Db	1966	CTCAGCGTGGCTGCCCCAAGGCCCTCAAGCTGCTTGTGCGCCGGTGGCTGCACC	2025
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCACTGACACCCACGATGCACTCAAGGAG	2040
Db	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCCTGCACTGACACCCACGATGCACTCAAGGAG	2085
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCACGGGCCTGCAC	2100
Db	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCACGGGCCTGCAC	2145
Qy	2101	TTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCGTGG	2160
Db	2146	TTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCCGTGG	2205
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAACGGAGATCCCCTTATCACATCTGAATGGCACG	2220
Db	2206	AAGAGTAAGCTCCTTGTCACTACCAACGGAGATCCCCTTATCACATCTGAATGGCACG	2265
Qy	2221	CAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGTGACCTG	2280
Db	2266	CAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGTGACCTG	2325

QY	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
QY	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2386	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2445
QY	2401	CCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2460
Db	2446	CCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2505
QY	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAGAAACTC	2520
Db	2506	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAGAAACTC	2565
QY	2521	CACCTGGACAGCCATCTCAGCTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db	2566	CACCTGGACAGCCATCTCAGCTCTTGCCCTCAAGCCCAGCCCCACAGCCATGATCCTC	2625
QY	2581	AACCTGTGGAGGCGCGGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Db	2626	AACCTGTGGAGGCGCGGCACTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2685
QY	2641	GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	2686	GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA	2742

RESULT 2

ADH71617

ID ADH71617 standard; DNA; 2752 BP.

XX

AC ADH71617;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21e SEQ ID NO:513.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71618.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 513; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2697; DB 12; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;

Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCGTGCCCAGTAGACGGCAGC	900
Db	886	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCGTGCCCAGTAGACGGCAGC	945
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
Qy	961	GAGTGCTCTGACCCAGCACCCCCAACGGAGGGAGGGAGTGCCAGGGACTGACCTGGAC	1020
Db	1006	GAGTGCTCTGACCCAGCACCCCCAACGGAGGGAGGGAGTGCCAGGGACTGACCTGGAC	1065
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1080
Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTGCC	1125
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGTCTGCTGCTGCTGTGCTCATC	1140
Db	1126	CTCTATGTGGCCTCATGCCGTGGCGCTGCCTGGTCTGCTGCTGCTGTGCTCATC	1185
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1186	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Qy	1201	ACCTCAGGCTTCCAGCCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1246	ACCTCAGGCTTCCAGCCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Qy	1261	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTGTCCCCGG	1320
Db	1306	ACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTGTCCCCGG	1365
Qy	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1366	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1425
Qy	1381	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGGCAGGAGTTCGTCTCC	1440
Db	1426	GGCGGCCGCCACACACTGCACCAACAGCTCTCCACCTCTGAGGGCAGGAGTTCGTCTCC	1485
Qy	1441	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1486	CGCCTCTCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1545
Qy	1501	GGGACCTTCAACTCCTGGGGCCGGTGTGATGATCCATAACAGGTATCAGCCTCCTC	1560
Db	1546	GGGACCTTCAACTCCTGGGGCCGGTGTGATGATCCATAACAGGTATCAGCCTCCTC	1605
Qy	1561	ATCCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1606	ATCCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCGTGCTGAGTCCATCGTTAGC	1680
Db	1666	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCGTGCTGAGTCCATCGTTAGC	1725

Qy	1681	TGTGGACCCCTGGCGTCTGTCACCCGGCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1726	TGTGGACCCCTGGCGTCTGTCACCCGGCAGTCATCCTGGCTATGGACCACTGTGGG	1785
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCAGGGCAGCTGG	1800
Db	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCAGGGCAGCTGG	1845
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTACTACTGCCAGCTGGAG	1860
Db	1846	GAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTACTACTGCCAGCTGGAG	1905
Qy	1861	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGAGAGGCC	1920
Db	1906	GCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGAGAGGCC	1965
Qy	1921	CTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	1980
Db	1966	CTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGGCCTGCACC	2025
Qy	1981	TCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	2026	TCCCTCGAGTACAACATCCGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
Qy	2041	GTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGGCCACGGGCCTGCAC	2100
Db	2086	GTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGGCCACGGGCCTGCAC	2145
Qy	2101	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2146	TTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2205
Qy	2161	AAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGAATGGCACG	2220
Db	2206	AAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGAATGGCACG	2265
Qy	2221	CAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2280
Db	2266	CAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGTGACCTG	2325
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2326	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATCAACTTC	2385
Qy	2341	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2386	AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2445
Qy	2401	CCAGCCCTGGTGGCCCCAGTCACCTCAAGATCCCTCCTCATTGGCAGAAGATAATT	2460
Db	2446	CCAGCCCTGGTGGCCCCAGTCACCTCAAGATCCCTCCTCATTGGCAGAAGATAATT	2505
Qy	2461	TCCAGCCTGGACCCACCCCTGTAAGCGGGGTGCCACTGGCGACTCTGGCCCAGAAACTC	2520
Db	2506	TCCAGCCTGGACCCACCCCTGTAAGCGGGGTGCCACTGGCGACTCTGGCCCAGAAACTC	2565
Qy	2521	CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGACCTC	2580

Db ||||||| 2566 CACCTGGACAGCCATCTCAGCTCTTGCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
Qy ||||||| 2581 AACCTGTGGAGGCAGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
Db ||||||| 2626 AACCTGTGGAGGCAGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685
Qy ||||||| 2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA 2697
Db ||||||| 2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGCTGA 2742

RESULT 3

ABK52891

ID ABK52891 standard; DNA; 2697 BP.

XX

AC ABK52891;

XX

DT 27-AUG-2002 (first entry)

XX

DE Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.

XX

KW Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW nootropic; neuroprotective; cytostatic; antiparkinsonian;
KW cerebroprotective; cancer; central nervous system; CNS; stroke;
KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .2697

FT /*tag= a

FT /product= "Netrin binding membrane receptor UNC5H-1"

XX

PN WO200233080-A2.

XX

PD 25-APR-2002.

XX

PF 15-OCT-2001; 2001WO-EP011891.

XX

PR 16-OCT-2000; 2000US-0240061P.

XX

PA (FARB) BAYER AG.

XX

PI Koehler RH;

XX

DR WPI; 2002-463314/49.

DR P-PSDB; AAU97899.

XX

PT Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.

XX

PS Claim 1; Fig 1; 94pp; English.

XX

CC This invention relates to the DNA and protein sequences of a novel

CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a nucleic
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC of the invention are useful to screen for agents which decrease the
CC activity of the UNC5H-1 protein. The sequences are also useful for
CC screening agents which regulate (modulate) the activity of the protein of
CC the invention. A pharmaceutical composition containing the protein of the
CC invention or a reagent that modulates the activity of the UNC5H-1 protein
CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g., Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents a DNA sequence encoding the
CC human netrin binding membrane receptor UNC5H-1 protein of the invention

XX

SQ Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Query Match 99.6%; Score 2687.4; DB 6; Length 2697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGGCCGTCCGGCCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 60
Db 1 ATGGCCGTCCGGCCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 60

Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120
Db 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAGCCA 180
Db 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAGCCA 180

Qy 181 GTGCTGCTTGTGTCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCAACGGGGAG 240
Db 181 GTGCTGCTTGTGTCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCAACGGGGAG 240

Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG 360
Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG 360

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAACGCC 420
Db 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAACGCC 420

Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db 421 TACATCCGCATAGCCATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy 481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGGCCAG 540

Db ||||||| 481 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGCATCCCTCAGCCGAG 540

Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATACTAC 600
|||

Db 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATACTAC 600

Qy 601 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCGCCTGCTGACACGCCAACTACACC 660
|||

Db 601 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCGCCTGCTGACACGCCAACTACACC 660

Qy 661 TGCCTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC 720
|||

Db 661 TGCCTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC 720

Qy 721 GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 780
|||

Db 721 GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 780

Qy 781 GGCTGGCAGAACCGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGTTTC 840
|||

Db 781 GGCTGGCAGAACCGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGTTTC 840

Qy 841 TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC 900
|||

Db 841 TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC 900

Qy 901 TGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT 960
|||

Db 901 TGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGAGCCGT 960

Qy 961 GAGTGCTCTGACCCAGCACCCCGAACGGAGGGGAGGAGTGCCAGGGACTGACCTGGAC 1020
|||

Db 961 GAGTGCTCTGACCCAGCACCCCGAACGGAGGGGAGGAGTGCCAGGGACTGACCTGGAC 1020

Qy 1021 ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTCTTGCCCTGAGGACGTGCC 1080
|||

Db 1021 ACCCGCAACTGTACCAAGTGACCTCTGTGTACACACTGCTTGCCCTGAGGACGTGCC 1080

Qy 1081 CTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGTGCTCATC 1140
|||

Db 1081 CTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGTGCTCATC 1140

Qy 1141 CTCGTTATTGCCGGAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
|||

Db 1141 CTCGTTATTGCCGGAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200

Qy 1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
|||

Db 1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260

Qy 1261 ACCATCCAGCCGGACCTCAGCACCAACCACCACTACAGGGCAGTCTGTCCCCGG 1320
|||

Db 1261 ACCATCCAGCCGGACCTCAGCACCAACCACCACTACAGGGCAGTCTGTCCCCGG 1320

Qy 1321 CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT 1380
|||

Db	1321	CAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTGGGT	1380
Qy	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTCGTCTCC	1440
Db	1381		
Db	1381	GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTCGTCTCC	1440
Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Db	1441		
Db	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACCTAT	1500
Qy	1501	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTCC	1560
Db	1501		
Db	1501	GGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAACAGGAATCAGCCTCC	1560
Qy	1561	ATCCCCCAGATGCCATACCCGAGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561		
Db	1561	ATCCCCCAGATGCCATACCCGAGGAAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621		
Db	1621	CCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Db	1681		
Db	1681	TGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCCTGTGGG	1740
Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741		
Db	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qy	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801		
Db	1801	GAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAGCTGGAG	1860
Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1920
Db	1861		
Db	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGTTGCCCTGGTGGAGAGGCC	1920
Qy	1921	CTCAGCGTGGCTGCCCCAAGGCCCTAACGCTGCTTGTGCGCCGGTGGCTGCACC	1980
Db	1921		
Db	1921	CTCAGCGTGGCTGCCCCAAGGCCCTAACGCTGCTTGTGCGCCGGTGGCTGCACC	1980
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981		
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCATGACACCCACGATGCACTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2100
Db	2041		
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Db	2101		
Db	2101	TTCAAGGACAGTTACCAACCTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGG	2160
Qy	2161	AAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAATGGCACG	2220
Db	2161		
Db	2161	AAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAATGGCACG	2220

Qy	2221 CAGCGGTACTTGCACCTCACCTGGAGCGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2221 CAGCGGTACTTGCACCTCACCTGGAGCGTCAGCCCCAGCACTAGTGACCTG	2280
Qy	2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Qy	2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Db	2341 AACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTC	2400
Qy	2401 CCAGCCCTGGTGGGCCAGTCAGCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2460
Db	2401 CCAGCCCTGGTGGGCCAGTCAGCTCAAGATCCCCTCCTCATTGGCAGAAGATAATT	2460
Qy	2461 TCCAGCCTGGACCCACCCTGTAGGCGGGTGCCGACTGGCGACTCTGGCCCAGAAACTC	2520
Db	2461 TCCAGCCTGGACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
Qy	2521 CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db	2521 CACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Qy	2581 AACCTGTGGAGGCGCGGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Db	2581 AACCTGTGGAGGCGCGGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2640
Qy	2641 GCTGGACTGGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGCTGA	2697
Db	2641 GCTGGACTGGGCCAGCCAGACGCTGGCTTTCACAGTGTGGAGGCTGAGTGCTGA	2697

RESULT 4

ABK49422

ID ABK49422 standard; DNA; 2881 BP.

XX

AC ABK49422;

XX

DT 15-JUL-2002 (first entry)

XX

DE DNA encoding human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
 KW cell signal processing; metabolic pathway modulation; cancerous tissue;
 KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
 KW chromosome 13; gene; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 87. .2786

FT /*tag= a

FT /product= "Human UNC5-like protein NOV1"

XX

PN WO200229038-A2.

XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031377.
XX
PR 04-OCT-2000; 2000US-0237862P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Herrmann JL, Rastelli L, Shimkets RA;
XX
DR WPI; 2002-340104/37.
DR P-PSDB; AAU79939.
XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT treating cardiomyopathy, atherosclerosis, and cancer.
XX
PS Claim 8; Page 7-8; 180pp; English.
XX
CC The present invention relates to a new NOVX polypeptide having a 900
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC residue amino acid sequence, as given in the specification. The novel
CC polypeptide, and its encoding polynucleotide, are used to treat
CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present nucleic acid sequence is that of the human UNC5-
CC like NOV1 gene located on chromosome 13. This sequence encodes the human
CC UNC5-like protein NOV1 of the invention
XX
SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2621.4; DB 6; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

Qy	1 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 60
Db	87 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 146
Qy	61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db	147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Qy	121 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
Db	207 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 266
Qy	181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGGAG 240
Db	267 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCAACGGGGAG 326
Qy	241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	327 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy	301	ACCATGGAGGTCCGCATTAATGTC TCAAGG CAGCAGG TCGAGAAGGTGTT CGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTC TCAAGG CAGCAGG TCGAGAAGGTGTT CGGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGC GTGGCATGGAGCT CCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGC GTGGCATGGAGCT CCTCGGGCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATT GCGCAAGAAC TT CGAGCAGGAGCCGCTGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATT GCGCAAGAAC TT CGAGCAGGAGCCGCTGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGC ATCGT GCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGC ATCGT GCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGGAA CGAGGACCTGGTGGACCCTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAA CGAGGACCTGGTGGACCCTGGACCCGTCCTGGACCCCAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGC CTGGTGGT GCGACAGGCCCGCCTTGCTGACACGCCA ACTACACC	660
Db	687	ACGCGGGAGCACAGC CTGGTGGT GCGACAGGCCCGCCTTGCTGACACGCCA ACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCGCAGCGCCTCCGCTGCTGT CATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCGCAGCGCCTCCGCTGCTGT CATCGTCTAC	806
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGTGGTCGACGTGGACCAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGGCCTCTCAACGGGGCGCTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGGCCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCTGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTGTCTGTGGACCGC	986
Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCA GTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCA GTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGCCTCATCGCCGTGCCGTCTGCCCTGGTCTGCTGCTGCTTGTCC	1137
Db	1167	GCCCTCTATGTGGCCTCATCGCCGTGCCGTCTGCCCTGGTCTGCTGCTGCTTGTCC	1226

Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCAACCACCACTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGGCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1437
Db	1464	GGTGGCGGGCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCTAATAACAGGTATCAGCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCTAATAACAGGTATCAGCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAACGCCCTCAAGCTGCTTGTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAACGCCCTCAAGCTGCTTGTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034

Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAACCTGGAGGAGATCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAACCTGGAGGAGATCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTCCTCATTGGCAGAACAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTCCTCATTGGCAGAACAG	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCGGGTGCCACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCGGGTGCCACTGGCGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCAGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCAGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC	2783
Qy	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 5

ADH71609

ID ADH71609 standard; DNA; 2881 BP.

XX

AC ADH71609;

XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV21a SEQ ID NO:505.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.

PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71610.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX
PS Example 21; SEQ ID NO 505; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2621.4; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Db 87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db 147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA 180
Db 207 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA 266

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCAACGGGGAG 240
Db 267 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCAACGGGGAG 326

Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 327 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db 387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 446

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 420
Db 447 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 506

Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db 507 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy	481	TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCACCGGAGGGATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCACCGGAGGGATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCAAAC TACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCAAAC TACACC	746
Qy	661	TGCGTGCCAAAGAACATCGTGGCACGTGCGCAGCGCCTCCGCTGCTGTATCGTCTAC	720
Db	747	TGCGTGCCAAAGAACATCGTGGCACGTGCGCAGCGCCTCCGCTGCTGTATCGTCTAC	806
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAACCGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTTC	840
Db	867	GGCTGGCAGAACCGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCTGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTGTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGGCCTCATGCCGTGGCGTCTGCCCTGCTGCTGCTGTCC	1137
Db	1167	GCCCTCTATGTGGGCCTCATGCCGTGGCGTCTGCCCTGCTGCTGCTGTCC	1226
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTCCAGCCGT CAGCATCAAGCCCAGCAAGCAGACAACCCCATCTG	1257
Db	1287	CTCACCTCAGGCTCCAGCCGT CAGCATCAAGCCCAGCAAGCAGACAACCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAAGGGCAGTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCACCAACCTACCAAGGGCAGTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377

Db	1404	CGGCAGGATGGGCCCAAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCCCTGGCTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCCCTGGCTCCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGGCCCTCAAGCTGCTCTGTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGGCCCTCAAGCTGCTCTGTTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTTACACATCTGGAAT	2214

Db	2244	CTGTGGAAGAGTAAGCTCCTGTCAGCTACCAGGAGATCCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGACGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGACGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAG	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCGGGTGCCACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCTGGCGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCCAAGCCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCGCCGACTTCCCAACGGCACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCGCCGACTTCCCAACGGCACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGCCAGCCAGCGCTGGCTCTTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGCCAGCCAGCGCTGGCTCTTCACAGTGTGGAGGCTGAGTGC	2783
Qy	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 6

ADH71633

ID ADH71633 standard; DNA; 2880 BP.

XX

AC ADH71633;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21m SEQ ID NO:529.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71634.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 529; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The present sequence encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2880 BP; 527 A; 984 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2880;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGCTTGGCTC 60
Db 86 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGCTTGGCTC 145

Qy 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCACACCG 120
Db 146 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCACACCG 205

Qy 121 GACCTGCTCCCCACTTCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
Db 206 GACCTGCTCCCCACTTCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 265

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACACGGGAG 240
Db 266 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACACGGGAG 325

Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 326 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 385

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db 386 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 445

Qy 361 GAATACTGGTGCAGTGCCTGGCATGGAGCTCCTGGCACCACCAAGAGTCAGAACGCC 420
Db 446 GAATACTGGTGCAGTGCCTGGCATGGAGCTCCTGGCACCACCAAGAGTCAGAACGCC 505

Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG 480
Db 506 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG 565

Qy 481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGAGGGCATCCCTCAGCCGAG 540
Db 566 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGAGGGCATCCCTCAGCCGAG 625

Qy 541 GTGGACTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db 626 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 685

Qy 601 ACAGCGGGAGCAGCAGCTGGTGGCGACAGGCCGCTTGCTGACACGGCCAACACACC 660
Db 686 ACAGCGGGAGCAGCAGCTGGTGGCGACAGGCCGCTTGCTGACACGGCCAACACACC 745

Qy 661 TGCCTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC 720

Db	746	TGC GTGG CCAAG AACA TCGT GGAC GAGT GGCT CGTGC AGCG CCCT CGTGT CATCG TCTAC	805
Qy	721	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC	780
Db	806	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC	865
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	840
Db	866	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	925
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCCACCCGTGCCAGTAGACGGC	897
Db	926	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGTCTGTGGACGGC	985
Qy	898	AGCTGGAGCCC GTGGAGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	986	AGCTGGAGCCC GTGGAGCAAGTGGTGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1045
Qy	958	CGTGAGT GCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTG	1017
Db	1046	CGTGAGT GCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTG	1105
Qy	1018	GACACCCGCAACTGTACCA GTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1106	GACACCCGCAACTGTACCA GTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1165
Qy	1078	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTTGTCC	1137
Db	1166	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTTGTCC	1225
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1226	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1285
Qy	1198	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAAACCCCCATCTG	1257
Db	1286	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAAACCCCCATCTG	1345
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCA ACCACCA CCTACCAGGGCAGTCTCTGTCCC	1317
Db	1346	CTCACCATCCAGCCGGACCTCAG---CACCA CCACCA CCTACCAGGGCAGTCTCTGTCCC	1402
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1403	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1462
Qy	1378	GGTGGCGGGCGCCACACACTGCACCA CAGCTCTCCACCTCTGAGGCCAGGGAGTTCGTC	1437
Db	1463	GGTGGCGGGCGCCACACACTGCACCA CAGCTCTCCACCTCTGAGGCCAGGGAGTTCGTC	1522
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAACATGACC	1497
Db	1523	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAACATGACC	1582
Qy	1498	TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557

Db	1583	TATGGGACCTCAACTCCTCGGGGCCGGCTGATGATCCATAACAGGTATCAGCCTC	1642
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1643	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1702
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1703	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGACCCCTGCTGAGTCCCATCGTT	1762
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1763	AGCTGTGGACCCCTGGCGTCTGCTACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1822
Qy	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1823	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGACCTCAAAAGCAGTCGTGCGAGGGCAGC	1882
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGAGGCCCTCCACCTCTACTACTGCCAG	1854
Db	1883	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGAGGCCCTCCACCTCTACTACTGCCAG	1942
Qy	1855	CTGGAGGCCAGTGCCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTTGCCTGGTGGGA	1914
Db	1943	CTGGAGGCCAGTGCCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTTGCCTGGTGGGA	2002
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTCTGTTGCCCGGTGCC	1974
Db	2003	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTCTGTTGCCCGGTGCC	2062
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCCTGCATGACACCCACGATGCACTC	2034
Db	2063	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCCTGCATGACACCCACGATGCACTC	2122
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2123	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC	2182
Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2183	CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2242
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAAT	2214
Db	2243	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAAT	2302
Qy	2215	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2303	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2362
Qy	2275	GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2334
Db	2363	GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2422
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2423	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2482

Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTCGCCTCAAGATCCCCTCCTCATCGGCAGAAG	2454
Db	2483	GGGGTCCCAGCCCTGGTGGGCCAGTCGCCTCAAGATCCCCTCCTCATCGGCAGAAG	2542
Qy	2455	ATAATTCCAGCCTGGACCCACCCGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAG	2514
Db	2543	ATAATTCCAGCCTGGACCCACCCGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCAG	2602
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2574
Db	2603	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2662
Qy	2575	ATCCTCAACCTGTGGAGGCAGCGCCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2663	ATCCTCAACCTGTGGAGGCAGCGCCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2722
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTTACAGTGTGGAGGCTGAGTGC	2694
Db	2723	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTTACAGTGTGGAGGCTGAGTGC	2782
Qy	2695	TGA 2697	
Db	2783	TGA 2785	

RESULT 7

ADH71649

ID ADH71649 standard; DNA; 2881 BP.

XX

AC ADH71649;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21u SEQ ID NO:545.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71650.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 545; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy

1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGCCTCGCCGCTTGGCTC 60

Db	87	ATGGCCGTCCGGCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGCTGGCTC 146
Qy	61	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGCCCAACCCAGTGCCTGGTGCCAACCCG 120
Db	147	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGCCCAACCCAGTGCCTGGTGCCAACCCG 206
Qy	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGAATGTGTACATCGTAAGAACAGCCA 180
Db	207	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGAATGTGTACATCGTAAGAACAGCCA 266
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCACGGGAG 240
Db	267	GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCACGGGAG 326
Qy	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	327	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 446
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCC 420
Db	447	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCC 506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG 480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG 566
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db	567	TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Qy	541	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db	627	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCAGACAGGCCGCTTGCTGACACGCCAACTACACC 660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCAGACAGGCCGCTTGCTGACACGCCAACTACACC 746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Qy	721	GTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 780
Db	807	GTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 866
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTTC 840
Db	867	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTTC 926
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCGTGCCAGTAGACGGC 897

Db 927 TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGCTTGTCTGTGGACGGC 986
Qy 898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCTGTGGACTGCACCCACTGGCGGAGC 957
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCTGTGGACTGCACCCACTGGCGGAGC 1046
Qy 958 CGTGAGTGCCTGACCCAGCACCCCGCAACGGAGGGAGGTGCCAGGGACTGACCTG 1017
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1047 CGTGAGTGCCTGACCCAGCACCCCGCAACGGAGGGAGGTGCCAGGGACTGACCTG 1106
Qy 1018 GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG 1077
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1107 GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG 1166
Qy 1078 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGCCTGCTGCTGCTTGTCCCTC 1137
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1167 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGCCTGCTGCTGCTTGTCCCTC 1226
Qy 1138 ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1227 ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Qy 1198 CTCACCTCAGGCTTCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1287 CTCACCTCAGGCTTCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Qy 1258 CTCACCATCCAGCCGGACCTCAGCACCAACCACCTACCAGGGCAGTCTGTCCC 1317
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCTACCAGGGCAGTCTGTCCC 1403
Qy 1318 CGGCAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG 1377
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1404 CGGCAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG 1463
Qy 1378 GGTGGCGGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGGAGTTCGTC 1437
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1464 GGTGGCGGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGGAGTTCGTC 1523
Qy 1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC 1497
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC 1583
Qy 1498 TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC 1557
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1584 TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCCTAACAGGTATCAGCCTC 1643
Qy 1558 CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC 1617
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1644 CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Qy 1618 AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCCTGCTGAGTCCCACCGTT 1677
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1704 AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCCTGCTGAGTCCCACCGTT 1763
Qy 1678 AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1764 AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCTCAAGCTGCTTCTGTTGCCTGGCGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCTCAAGCTGCTTCTGTTGCCTGGCGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGTCCCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCTCATTGGCAGAAC	2454
Db	2484	GGGTCCCAGCCCTGGTGGCCCCAGTGCCTCAAGATCCCTCATTGGCAGAAC	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTGCCTCCAAGCCCAGCCCCACAGCCATG	2663

QY 2575 ATCCTAACCTGTGGGAGGCGCGGCAC TCCCCAACGGAACCTCAGCCAGCTGGCTGCA 2634
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2664 ATCCTAACCTGCGGGAGGCGCGGCAC TCCCCAACGGAACCTCAGCCAGCTGGCTGCA 2723

QY 2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC 2694
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC 2783

QY 2695 TGA 2697
| | | .
Db 2784 TGA 2786

RESULT 8

ADH71635

ID ADH71635 standard; DNA; 2881 BP.

XX

AC ADH71635;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21n SEQ ID NO:531.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

xx

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

xx

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71636.

xx

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

xx

PS Example 21; SEQ ID NO 531; 1880pp; English.

xx

The invention relates to a novel isolated polypeptide (NOVX). A polypeptide of the invention has cytostatic, immunomodulator, neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and antilipidaemic activity, and may have a use in gene therapy, and as a vaccine. The polypeptides are encoded by NOVX polynucleotides comprising any of the 303 fully defined nucleotide sequences given in the specification. The polypeptide is useful in the manufacture of a medicament for treating a syndrome associated with a human disease. The polypeptide, polynucleotide and antibody are useful in diagnosing, treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The present sequence encodes a NOVX polypeptide of the invention.

xx

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

QV

1 ATGGCCGTCCGGCCCCGGCTGTGGCCAGCGCTCTGGGCATAGTCCTCGCCGTTGGCTC 60

Ph

87 ATGGCCGTCCGGCCCAGCCTGTGGCCAGCGCTCTGGCATAGTCCTCGCCGTTGGCTC 146
61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCAACCCG 120
|||||||||||||||||||||||||||||||||||||||||||

121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180

1

181 GTGCTGTTGTGTCAAGGCCGTCCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Db	267	GTGCTGCTGTGTCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCAACGGGGAG	326
Qy	241	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGCTGCC	300
Db	327	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGAGCAGTGGTGAGCCG	386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGAAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	686.
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCCAACATACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCCAACATACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	806
Qy	721	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGTTTC	840
Db	867	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGTTTC	926
Qy	841	TGTGAGGGCAGAACATGTCCAGAA---AACAGCCTGCCACCCGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAACATGTCCATGACCGCACCGTCTCTCTGCTTGTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGTGGCCTGTGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTGGCCTGTGGCTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG	1017
Db	1047	CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG	1166

Qy	1078	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCTGCTGCTGCTGCCTC 1137
Db	1167	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCTGCTGCTGCTGCCTC 1226
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCATT 1197
Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCATT 1286
Qy	1198	CTCACCTCAGGCTTCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
Db	1287	CTCACCTCAGGCTTCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCC 1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCACCACTACCAGGGCAGTCTCTGTCCC 1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG 1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG 1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC 1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC 1583
Qy	1498	TATGGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
Db	1584	TATGGGACCTTCAACTTCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT 1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT 1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC 1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC 1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG 1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG 1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGGA 1914
Db	1944	CTGGAGGCCAGTGCCTGCTACATCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGGA 2003

Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCCGCCGGTGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCCGCCGGTGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2543
Qy	2455	ATAATTCCAGCCTGGACACCACCCACTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACACCACCCACTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCGCGGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCGCGGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC	2783
Qy	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 9
ADH71637
ID ADH71637 standard; DNA; 2881 BP.
XX
AC ADH71637;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV21o SEQ ID NO:533.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71638.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 533; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Db 87 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db 147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
Db 207 GACCTGCTCCCCACTTCCTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 266

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCACGGGGAG 240
Db 267 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTTCTCAAGTGCACGGGGAG 326

Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 327 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAG 360
Db 387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAG 446

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Db 447 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACTC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACTC	686
Qy	601	ACGCAGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCCAACCTACACC	660
Db	687	ACGCAGGGAGCACAGCCTGGTGGTGCACAGGCCGCCCTGCTGACACGGCCAACCTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	806
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCCACCCCTGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTTGTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCCTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCCTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCAGGGACTGACCTG	1017
Db	1047	CGTGAGTGCCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCAGGGACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAAGTACCTCTGTGTACACAGTGCCTCTGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAAGTACCTCTGTGTACACAGTGCCTCTGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTGTCC	1137
Db	1167	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGTCTGCTGCTGCTGTCC	1226
Qy	1138	ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346

Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCACCAACCTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGGCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1437
Db	1464	GGTGGCGGGCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTCACCGAGCAGCTGGCCGTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAAGCTGCTTGTGCGCCGGTGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2154

Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAAGGAGATCCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACCTTCACCCCTGGAGCGTGTCAAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCGCGGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCGCGGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTGGAGGCTGAGTGC	2783
Qy	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 10

ADH71641

ID ADH71641 standard; DNA; 2881 BP.

XX

AC ADH71641;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21q SEQ ID NO:537.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71642.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 537; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the specification. The polypeptide is useful in the manufacture of a medicament for treating a syndrome associated with a human disease. The polypeptide, polynucleotide and antibody are useful in diagnosing, treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The present sequence encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 986 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 60
Db 87 ATGGCCGTCCGGCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db 147 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
Db 207 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 266

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCAACGGGAG 240
Db 267 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCAACGGGAG 326

Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 327 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
Db 387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 446

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 420
Db 447 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 506

Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db 507 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy 481 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db 567 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGACCCGTCCTGGACCCCAATGTATACATC 600
Db 627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGACCCGTCCTGGACCCCAATGTATACATC 686

Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCGCCTGCTGACACGGCAACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	806
Qy	721	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCCACCCGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGCTTGTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCCTGGAGCAAGTGGCGGCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCCTGGAGCAAGTGGCGGCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCCTCTGACCCAGCACCCCGAACGGAGGGAGGTGCCAGGGACTGACCTG	1017
Db	1047	CGTGAGTGCCTCTGACCCAGCACCCCGAACGGAGGGAGGTGCCAGGGACTGACCTG	1106
Qy	1018	GACACCCGAACTGTACCAAGTACCTCTGTGTACACAGTGCCTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGAACTGTACCAAGTACCTCTGTGTACACAGTGCCTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCTGCTGCTGCTGTCTC	1137
Db	1167	GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCTGCTGCTGCTGTCTC	1226
Qy	1138	ATCCTCGTTATTGCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGACCTCAGCACCAACCACCTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGACCTCAG---CACCAACCACCTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCAGGGAGTCGTC	1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCAGGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1497

Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTCGGGGCGCGTGTGATGATCCCTAACATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGCGCGTGTGATGATCCCTAACATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAACGCGCTCAAGCTGCTTCTGTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAACGCGCTCAAGCTGCTTCTGTTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGCTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCTTGTAGCTACCAGGAGATCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCTTGTAGCTACCAGGAGATCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCAGTCACCTCACCCCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCAGTCACCTCACCCCTGGAGCGTGTAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGCGGAGCTTCAGCATC	2334

Db 2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
 Qy 2335 AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
 |||||||
 Db 2424 AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
 |||||||
 Qy 2395 GGGTCCCAGCCCTGGTGGGCCAGTCAGCTCAAGATCCCCTCCTCATTGGCAGAAC 2454
 |||||||
 Db 2484 GGGTCCCAGCCCTGGTGGGCCAGTCAGCTCAAGATCCCCTCCTCATTGGCAGAAC 2543
 |||||||
 Qy 2455 ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG 2514
 |||||||
 Db 2544 ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG 2603
 |||||||
 Qy 2515 AAACCTCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAGCCATG 2574
 |||||||
 Db 2604 AAACCTCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGGCCAGCCCCACAGCCATG 2663
 |||||||
 Qy 2575 ATCCTCACCTGTGGAGGCGCCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
 |||||||
 Db 2664 ATCCTCACCTGTGGAGGCGCCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723
 |||||||
 Qy 2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2694
 |||||||
 Db 2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2783
 |||||||
 Qy 2695 TGA 2697
 |||
 Db 2784 TGA 2786

RESULT 11

ADH71629

ID ADH71629 standard; DNA; 2881 BP.

XX

AC ADH71629;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21k SEQ ID NO:525.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.

PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71630.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 525; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy	1 ATGGCCGTCCGGCCCCGGCTGTGGCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Db	87 ATGGCCGTCCGGCCCCGGCTGTGGCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 146
Qy	61 CGCGGCTCGGGTGCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCAACCCG 120
Db	147 CGCGGCTCGGGTGCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCAACCCG 206
Qy	121 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
Db	207 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 266
Qy	181 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCAACGGGAG 240
Db	267 GTGCTGCTTGTGTGCAAGGCCGTGCCGACGCAGATCTCTCAAGTGCAACGGGAG 326
Qy	241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db	327 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Qy	301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAG 360
Db	387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGGCTGGAG 446
Qy	361 GAATACTGGTGCAGTGCCTGGCATGGAGCTCCTGGGACCCACCAAGAGTCAGAAGGCC 420
Db	447 GAATACTGGTGCAGTGCCTGGCATGGAGCTCCTGGGACCCACCAAGAGTCAGAAGGCC 506
Qy	421 TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCCTGGCCAAGGAGGTG 480
Db	507 TACATCCGCATAGCCAGATTGCGCAAGAACATTGAGCAGGAGCCCTGGCCAAGGAGGTG 566
Qy	481 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db	567 TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
Qy	541 GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db	627 GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
Qy	601 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCCGCCTGCTGACACGGCCAACATACACC 660
Db	687 ACGCGGGAGCACAGCCTGGTGGCGACAGGCCCGCCTGCTGACACGGCCAACATACACC 746
Qy	661 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTACATCGTCTAC 720
Db	747 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTACATCGTCTAC 806
Qy	721 GTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 780
Db	807 GTGAACGGTGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGC 866
Qy	781 GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTTC 840

Db	867	GGCTGGCAGAACCGGAGCCGGAGCTGCACCAACC CGGCCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCGTGCCCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTTGCTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCC GTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCC GTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCATGTGGCCTCATGCCGTGCCGTCTGCCTGGCCTGCTGCTGCTTGTCCCTC	1137
Db	1167	GCCCTCATGTGGCCTCATGCCGTGCCGTCTGCCTGGCCTGCTGCTGCTTGTCCCC	1226
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCGT CAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCGT CAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCGGACCTCAGCACCAACCACCTACCAAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCGGACCTCAG---CACCACCAACCTACCAAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGTGTCAAGACCCCTGCTGAGTCCCACCGT	1677

Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAAGACCCGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCCTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTCTGTTGCCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCCTCAAGCTGCTTCTGTTGCCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTAAGGACAGTTACCACAAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAAGCTACCACTGGAGATCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAAGCTACCACTGGAGATCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACTGCACCTCACCTGGAGCGTGTAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCAGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTCCTCATCGGCAGAAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTCCTCATCGGCAGAAG	2543
Qy	2455	ATAATTTCAGCCTGGACCCACCCACCTGTAGGGGGGTGCCGACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTTCAGCCTGGACCCACCCACCTGTAGGGGGGTGCCGACTGGCGACTCTGGCCAG	2603

Qy 2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGCCATG 2574
|||
Db 2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy 2575 ATCCTCAACCTGTGGAGGCGCGGACTTCCCAACGGAACCTCAGCCAGCTGGCTGCA 2634
|||
Db 2664 ATCCTCAACCTGTGGAGGCGCGGACTTCCCAACGGAACCTCAGCCAGCTGGCTGCA 2723

Qy 2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGC 2694
|||
Db 2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGC 2783

Qy 2695 TGA 2697
|||
Db 2784 TGA 2786

RESULT 12

ADH71631

ID ADH71631 standard; DNA; 2881 BP.

XX

AC ADH71631;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV211 SEQ ID NO:527.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71632.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 527; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 526 A; 984 C; 868 G; 503 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGCTTGGCTC 60
Db |||||||

Db 87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGCTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120
Db |||||||

Db 147 CGCGGCTCGGGTGCCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180

Db	207	GACCTGCTTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA	266
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGAACGGGGAG	240
Db	267	GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGAACGGGGAG	326
Qy	241	TGGGTGCCGCCAGGTGGACCACGTGATCGAGCCAGCACAGACGGGAGCAGTGGCTGCC	300
Db	327	TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG	386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCTTGCTGACACGGCCAACATACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCTTGCTGACACGGCCAACATACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTGCAGCGCTCCGCTGCTGTATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTGCAGCGCTCCGCTGCTGTATCGTCTAC	806
Qy	721	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	807	GTGAACGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	866
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGCGCTTC	926
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCAGGCCACCTGTGCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCTCTGCTTGTCTGTGGACCGC	986
Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGCGCTGTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGCGCTGTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG	1017

Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCCTGCTGCTGCTGTCCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATGCCGTGGCGTCTGCCTGGTCCTGCTGCTGCTGTCCCTC	1226
Qy	1138	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCACACCACCTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCACCAACCTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTCGTC	1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGTCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCCCCTCCACCTCTACTACTGCCAG	1943

Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGCCGCTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACCGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT	2274
Db	2304	GGCACCGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAACAG	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCGGGTGCCGACTGGCGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTTACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTTACAGTGTGGAGGCTGAGTGC	2783

QY 2695 TGA 2697

|||

Db 2784 TGA 2786

RESULT 13

ADH71645

ID ADH71645 standard; DNA; 2881 BP.

XX

AC ADH71645;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21s SEQ ID NO:541.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;

PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71646.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 541; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 60
Db 87 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGCATAGTCCTGCCGTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120
Db 147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 180
Db 207 GACCTGCTCCCCACTTCCTGGTGAGCCGAGGATGTGTACATCGTCAAGAACAAAGCCA 266

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 240
Db 267 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGGAG 326

Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 327 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360

Db 387 ACCATGGAGGTCCGCATTAATGTCCTCAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG 446
Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGACCACCAAGAGTCAGAAGGCC 420
|||
Db 447 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGACCACCAAGAGTCAGAAGGCC 506
Qy 421 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
|||
Db 507 TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
Qy 481 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG 540
|||
Db 567 TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG 626
Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC 600
|||
Db 627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC 686
Qy 601 ACGCGGGAGCACAGCCTGGTGGTGCAGACAGGCCGCTTGCTGACACGGCCAACATACACC 660
|||
Db 687 ACGCGGGAGCACAGCCTGGTGGTGCAGACAGGCCGCTTGCTGACACGGCCAACATACACC 746
Qy 661 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC 720
|||
Db 747 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGGCCCTCCGCTGCTGTATCGTCTAC 806
Qy 721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 780
|||
Db 807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 866
Qy 781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCGGCCCTCTAACGGGGCGTTTC 840
|||
Db 867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCGGCCCTCTAACGGGGCGTTTC 926
Qy 841 TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCCACCCCTGTGCCAGTAGACGGC 897
|||
Db 927 TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGCTTGTCTGTGGACGGC 986
Qy 898 AGCTGGAGCCGTGGAGCAAGTGGTCGCCCTGTGGACTGCACCCACTGGCGGAGC 957
|||
Db 987 AGCTGGAGCCGTGGAGCAAGTGGTCGCCCTGTGGACTGCACCCACTGGCGGAGC 1046
Qy 958 CGTGAGTGCCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG 1017
|||
Db 1047 CGTGAGTGCCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCAGGGACTGACCTG 1106
Qy 1018 GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTGTGCCCTGAGGACGTG 1077
|||
Db 1107 GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTTGTGCCCTGAGGACGTG 1166
Qy 1078 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGCTGTGCTGCTGTCTC 1137
|||
Db 1167 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGCTGTGCTGCTGTCTC 1226
Qy 1138 ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
|||
Db 1227 ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy	1198	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAG---CACCAACCACCACTACCAGGGCAGTCTCTGTCCC	1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTACCAATGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTTGAGGCCAGGAGTCGTC	1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCACCTTGAGGCCAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATAACCCGAGGGAAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCCGTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGGCCCTCAAGCTGCTTCTGTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGGCCCTCAAGCTGCTTCTGTTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123

QY	2035	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
QY	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCAGCTCC	2243
QY	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAGGAGATCCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACCAGGAGATCCCCTTTATCACATCTGGAAT	2303
QY	2215	GGCACGCAGCGGTACTTGCACGCACCTTCACCCCTGGAGCGTGTAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACCGCACCTTCACCCCTGGAGCGTGTAGCCCCAGCACTAGT	2363
QY	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTCAGCATC	2423
QY	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
QY	2395	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTCCTCATTGGCAGAAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTCCTCATTGGCAGAAG	2543
QY	2455	ATAATTCCAGCCTGGACACCCACCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACACCCACCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2603
QY	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCTCCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCTCCAAGCCCAGCCCCACAGCCATG	2663
QY	2575	ATCCTCAACCTGTGGAGGCAGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCAGCGGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
QY	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTCACAGTGTGGAGGCTGAGTGC	2783
QY	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 14

ADH71627

ID ADH71627 standard; DNA; 2881 BP.

XX

AC ADH71627;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21j SEQ ID NO:523.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX
PA (CURA-) CURAGEN CORP.
XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71628.

XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX
PS Example 21; SEQ ID NO 523; 1880pp; English.
XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy	1	ATGGCCGTCCGGCCCGGCCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC	60
Db	87	ATGGCCGTCCGGCCCGGCCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC	146
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCACACCG	120
Db	147	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCACACCG	206
Qy	121	GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA	180
Db	207	GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTAAGAACAAAGCCA	266
Qy	181	GTGCTGCTTGTGTCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACACGGGAG	240
Db	267	GTGCTGCTTGTGTCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACACGGGAG	326
Qy	241	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC	300
Db	327	TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCC	386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626

Qy 541 GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATAACATC 600
Db 627 GTGGAGTGGCTCCGAACGAGGACCTGGTGGACCCGTCCCTGGACCCAATGTATAACATC 686
Qy 601 ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCTGCTGACACGCCAACTACACC 660
Db 687 ACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCTGCTGACACGCCAACTACACC 746
Qy 661 TGCCTGGCCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
Db 747 TGCCTGGCCAAGAACATCGTGGCACGTCGCCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
Qy 721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 780
Db 807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC 866
Qy 781 GGCTGGCAGAACGGAGGCCAGCTGCACCAACCCGGGCCCTCTCAACGGGGCGTTTC 840
Db 867 GGCTGGCAGAACGGAGGCCAGCTGCACCAACCCGGGCCCTCTCAACGGGGCGTTTC 926
Qy 841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCCACCCGTGCCCCAGTAGACGGC 897
Db 927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCTCTGTGCTGTCTGTGGACGGC 986
Qy 898 AGCTGGAGCCGTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC 957
Db 987 AGCTGGAGCCGTGGAGCAAGTGGTGGCCTGTGGACTGCACCCACTGGCGGAGC 1046
Qy 958 CGTGAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGTGCCAGGGCACTGACCTG 1017
Db 1047 CGTGAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGTGCCAGGGCACTGACCTG 1106
Qy 1018 GACACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCTCTGGCCCTGAGGACGTG 1077
Db 1107 GACACCCGCAACTGTACCAAGTGACCTCTGTGTGACAGTGCTCTGGCCCTGAGGACGTG 1166
Qy 1078 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTGTCCCTC 1137
Db 1167 GCCCTCTATGTGGCCTCATGCCGTGGCGTCTGCCTGGCCTGCTGCTGCTGTCCCTC 1226
Qy 1138 ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
Db 1227 ATCCTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Qy 1198 CTCACCTCAGGCTTCCAGCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
Db 1287 CTCACCTCAGGCTTCCAGCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Qy 1258 CTCACCATCCAGCCGGACCTCAGCACCAACCACCACTACCAGGGCAGTCTCTGTCCC 1317
Db 1347 CTCACCATCCAGCCGGACCTCAG---CACCACCAACCACTACCAGGGCAGTCTCTGTCCC 1403
Qy 1318 CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
Db 1404 CGGCAGGATGGGCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy	1378	GGTGGCGGCCGACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1437
Db	1464	GGTGGCGGCCGACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGGCCTCAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCCCTCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTCCTGCTACGTCTTCACCGAGCAGCTGGCCCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTCCTGCTACGTCTTCACCGAGCAGCTGGCCCTTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGCTGCTTGTGTTGCCCGGTGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGCTGCTTGTGTTGCCCGGTGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGCTACTGCCTGCATGACACCCACGATGCACTC	2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAACAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCAACCTGCGCTATCCATCCACGATGTGCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCACTACAGGAGATCCCCTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGGTACTTGCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT	2274

Db	2304	GGCACGCAGCGGTACTTGCACGCCCTGAGCGTGTCAAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTTCCTCATTGGCAGAACAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCTTCCTCATTGGCAGAACAG	2543
Qy	2455	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2514
Db	2544	ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGGACTCTGGCCAG	2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTGCCCTCAAGCCCAGCCCCACAGCCATG	2663
Qy	2575	ATCCTCAACCTGTGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
Db	2664	ATCCTCAACCTGTGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTTTCACAGTGTGGAGGCTGAGTGC	2783
Qy	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 15

ADH71639

ID ADH71639 standard; DNA; 2881 BP.

xx

AC ADH71639;

xx

DT 25-MAR-2004 (first entry)

xx

DE Human gene of the invention NOV21p SEQ ID NO:535.

xx

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

xx

OS *Homo sapiens.*

xx

PN WO2003102155-A2.

xx

PD 11-DEC-2003.

XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Paturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR P-PSDB; ADH71640.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 535; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 60
Db 87 ATGGCCGTCCGGCCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGTTGGCTC 146

Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120
Db 147 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 206

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA 180
Db 207 GACCTGCTCCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA 266

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACGGGAG 240
Db 267 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTTCTCAAGTGCACGGGAG 326

Qy 241 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
Db 327 TGGGTGCCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG 360
Db 387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTGGCTGGAG 446

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 420
Db 447 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAACGCC 506

Qy 421 TACATCCGCATAGCCAGATTGCCAAGAACATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db 507 TACATCCGCATAGCCAGATTGCCAAGAACATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy 481 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCGAG 540
Db 567 TCCCTGGAGCAGGGCATCGTGTGCCCTGCCGTCCACCGGAGGGCATCCCTCAGCCGAG 626

Qy 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db 627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy 601 ACGCGGGAGCACAGCCTGGTGGTGCAGAGGCCGCTTGCTGACACGGCCAACATACACC 660
Db 687 ACGCGGGAGCACAGCCTGGTGGTGCAGAGGCCGCTTGCTGACACGGCCAACATACACC 746

Qy 661 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGCGCCTCCGCTGCTGTATCGTCTAC 720
Db 747 TGCCTGGCCAAGAACATCGTGGCACGTGCCGCAGCGCCTCCGCTGCTGTATCGTCTAC 806

Qy	721	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC 780
Db	807	GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGCTGCAGGCCAGCTGTGGCGC 866
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCACACGGGGCGCTTC 840
Db	867	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCACACGGGGCGCTTC 926
Qy	841	TGTGAGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCGTGCCCCAGTAGACGGC 897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGCTTGCTCTGTGGACGGC 986
Qy	898	AGCTGGAGCCC GTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
Db	987	AGCTGGAGCCC GTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCAGGGCACTGACCTG 1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCAGGGCACTGACCTG 1106
Qy	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG 1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGCCCTGAGGACGTG 1166
Qy	1078	GCCCTCTATGTGGGCCTCATGCCGTGGCCGTCTGCCCTGGCTCTGCTGCTGCTTGTCCC 1137
Db	1167	GCCCTCTATGTGGGCCTCATGCCGTGGCCGTCTGCCCTGGCTCTGCTGCTGCTTGTCCC 1226
Qy	1138	ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
Db	1227	ATCCTCGTTATTGCCCGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Qy	1198	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
Db	1287	CTCACCTCAGGCTTCCAGCCCCGTCAAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
Qy	1258	CTCACCATCCAGCCGACCTCAGCACCAACCACCTACCAGGGCAGTCTCTGTCCC 1317
Db	1347	CTCACCATCCAGCCGACCTCAG---CACCAACCACCTACCAGGGCAGTCTCTGTCCC 1403
Qy	1318	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG 1377
Db	1404	CGGCAGGATGGGCCAGCCCCAAGTCCAGCTCACCAATGGCACCTGCTCAGCCCCCTG 1463
Qy	1378	GGTGGCGGCCGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1437
Db	1464	GGTGGCGGCCGCCACACACTGCACCAAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC 1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAGCAACATGACC 1583
Qy	1498	TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCATAACAGGTATCAGCCTC 1557
Db	1584	TATGGGACCTTCAACTTCCCTGGGGCCGGCTGATGATCCATAACAGGTATCAGCCTC 1643
Qy	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617

Db	1644	CTCATCCCCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703
Qy	1618	AAGCCGGAAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT 1677
Db	1704	AAGCCGGAAAGACGTGAGGTTGCCCTAGCTGGCTGTCAGACCCCTGCTGAGTCCCATCGTT 1763
Qy	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACGT 1737
Db	1764	AGCTGTGGACCCCTGGCGTCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACGT 1823
Qy	1738	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
Db	1824	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883
Qy	1798	TGGG---AGGATGTGCTGCACCTGGCGAGGAGGCAGCCCTCCACCTCTACTACTGCCAG 1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGCGAGGAGGCAGCCCTCCACCTCTACTACTGCCAG 1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGGA 1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGCCGCTTGCCCTGGTGGGA 2003
Qy	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTGTGTTGCCCGGTGCC 1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAGGCTGCTTGTGTTGCCCGGTGCC 2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCTGCATGACACCCACGATGCACTC 2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGCTACTGCCTGCATGACACCCACGATGCACTC 2123
Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC 2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTC 2183
Qy	2095	CTGCACTTCAAGGACAGTTACCAACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2154
Db	2184	CTGCACTTCAAGGACAGTTACCAACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAAGTACCAAGGAGATCCCTTTATCACATCTGGAAT 2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAAGTACCAAGGAGATCCCTTTATCACATCTGGAAT 2303
Qy	2215	GGCACGCAGCGGTACTTGCAGTCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT 2274
Db	2304	GGCACGCAGCGGTACTTGCAGTCACCTCACCTGGAGCGTGTCAAGCCCCAGCACTAGT 2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGAGGGCAGGGCAGAGCTTCAGCATC 2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGAGGGCAGAGCTTCAGCATC 2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCAGTGCCTCAAGATCCCCTCCTCATTGGCAGAAG 2454

Db 2484 GGGTCCCAGCCCTGGTGGGCCCAAGTCCTCAAGATCCCCTCCTCATCGGCAGAAG 2543
Qy 2455 ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG 2514
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2544 ATAATTCCAGCCTGGACCCACCCCTGTAGGCAGGGTGCCGACTGGCGACTCTGGCCAG 2603
Qy 2515 AAACCTCACCTGGACAGCCATCTCAGCTTCTTGCCCTCCAAGCCCAGCCCCACAGCCATG 2574
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2604 AAACCTCACCTGGACAGCCATCTCAGCTTCTTGCCCTCCAAGCCCAGCCCCACAGCCATG 2663
Qy 2575 ATCCTCAACCTGTGGAGGCGGGCACTTCCCCAACGGAACCTCAGCCAGCTGGCTGCA 2634
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2664 ATCCTCAACCTGTGGAGGCGGGCACTTCCCCAACGGAACCTCAGCCAGCTGGCTGCA 2723
Qy 2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2694
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTCACAGTGTGGAGGCTGAGTGC 2783
Qy 2695 TGA 2697
| | |
Db 2784 TGA 2786

Search completed: March 5, 2005, 22:58:58
Job time : 1381 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 19:26:23 ; Search time 424.67 Seconds
(without alignments)
10391.692 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PECTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			Description	
		Match	Length	DB	ID	
1	2252.2	83.5	3014	2	US-08-808-982-1	Sequence 1, Appli
2	2252.2	83.5	3014	3	US-09-306-902A-1	Sequence 1, Appli
3	1552.4	57.6	1787	2	US-08-808-982-2	Sequence 2, Appli
4	1552.4	57.6	1787	3	US-09-306-902A-2	Sequence 2, Appli
5	841.4	31.2	2831	2	US-08-808-982-3	Sequence 3, Appli
6	841.4	31.2	2831	3	US-09-306-902A-3	Sequence 3, Appli
7	833.6	30.9	3008	4	US-09-949-016-4794	Sequence 4794, Ap
8	487	18.1	2736	4	US-09-969-532-9	Sequence 9, Appli
9	487	18.1	3411	4	US-09-969-532-33	Sequence 33, Appli
10	467	17.3	2703	4	US-09-969-532-11	Sequence 11, Appli
11	460.2	17.1	2694	4	US-09-969-532-13	Sequence 13, Appli

12	439.8	16.3	2661	4	US-09-969-532-15	Sequence 15, Appl	
13	323	12.0	349	4	US-09-471-276-345	Sequence 345, App	
14	293.2	10.9	1968	4	US-09-969-532-31	Sequence 31, Appl	
15	293.2	10.9	2001	4	US-09-969-532-29	Sequence 29, Appl	
16	293.2	10.9	2010	4	US-09-969-532-27	Sequence 27, Appl	
17	293.2	10.9	2043	4	US-09-969-532-25	Sequence 25, Appl	
18	284.8	10.6	1659	4	US-09-969-532-7	Sequence 7, Appli	
19	284.8	10.6	1692	4	US-09-969-532-5	Sequence 5, Appli	
20	284.8	10.6	1701	4	US-09-969-532-3	Sequence 3, Appli	
21	284.8	10.6	1734	4	US-09-969-532-1	Sequence 1, Appli	
c	22	269	10.0	771	1	US-08-253-155A-17	Sequence 17, Appl
c	23	181.4	6.7	19326	4	US-09-949-016-16776	Sequence 16776, A
	24	123.4	4.6	114139	4	US-09-949-016-16536	Sequence 16536, A
	25	119	4.4	305	2	US-08-808-982-4	Sequence 4, Appli
	26	119	4.4	305	3	US-09-306-902A-4	Sequence 4, Appli
c	27	80.4	3.0	601	4	US-09-949-016-169999	Sequence 169999,
	28	80.2	3.0	966	4	US-09-969-532-23	Sequence 23, Appl
	29	75.8	2.8	1008	4	US-09-969-532-19	Sequence 19, Appl
c	30	73.8	2.7	601	4	US-09-949-016-170033	Sequence 170033,
c	31	66.4	2.5	601	4	US-09-949-016-170040	Sequence 170040,
c	32	66.4	2.5	601	4	US-09-949-016-170041	Sequence 170041,
	33	63.6	2.4	5784	4	US-09-949-016-462	Sequence 462, App
c	34	63	2.3	601	4	US-09-949-016-170000	Sequence 170000,
c	35	63	2.3	601	4	US-09-949-016-170034	Sequence 170034,
	36	60.2	2.2	2820	4	US-09-854-845-15	Sequence 15, Appl
	37	60.2	2.2	2865	4	US-09-854-845-13	Sequence 13, Appl
	38	60.2	2.2	3105	4	US-09-854-845-5	Sequence 5, Appli
	39	60.2	2.2	3150	4	US-09-854-845-1	Sequence 1, Appli
	40	60.2	2.2	3237	4	US-09-854-845-7	Sequence 7, Appli
	41	60.2	2.2	3282	4	US-09-854-845-3	Sequence 3, Appli
	42	60.2	2.2	3411	4	US-09-854-845-11	Sequence 11, Appl
	43	60.2	2.2	3456	4	US-09-854-845-9	Sequence 9, Appli
	44	60.2	2.2	4074	4	US-09-854-845-17	Sequence 17, Appl
	45	56.4	2.1	5178	4	US-09-949-016-5241	Sequence 5241, Ap

ALIGNMENTS

RESULT 1
 US-08-808-982-1
 ; Sequence 1, Application US/08808982
 ; Patent No. 5939271
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; APPLICANT: Leonardo, E. David
 ; APPLICANT: Hink, Lindsay
 ; APPLICANT: Masu, Masayuki
 ; APPLICANT: Kazuko, Keino-Masu
 ; TITLE OF INVENTION: Netrin Receptors
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA

; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

US-08-808-982-1

Query Match 83.5%; Score 2252.2; DB 2; Length 3014;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

Qy 1 ATGGCCGTCCGGCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTGCCGTTGGCTC 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 ATGGCCGTCCGGCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTGCCGCTGGCTT 60

Qy 61 CGCGGCTCGGGTGCCAGCAGAGTGCCACCGTGGCAACCCAGTGCCTGGTGCCAACCCG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CGTGGTTGGGTGCCAGCAGAGTGCCACGGTGGCAATCCAGTGCCCGGTGCCAACCC 120

Qy 121 GACCTGCTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTAAGAACAAAGCCA 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GACCTGCTGCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAAGCCG 180

Qy 181 GTGCTGCTTGTGTGCAAGGCCGTGCCGCCACGCAGATCTCTCAAGTGCACGGGAG 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTCTCAAGTGCACGGGAA 240

Qy 241 TGGGTGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGCTGCC 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TGGGTGCCAGGTGATCACGTAATTGAACGCAGCACCGACAGCAGCGGGATTGCCA 300

Qy 301 ACCATGGAGGTCCGCATTAATGTCCTCAAGGCAGCAGGTGAGAAGGTGTTGGCTGGAG 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 ACCATGGAGGTCCGTATCACGTATCGAGGCAGCAGGTAGAGAAAGTGTGGCTGGAG 360

Qy 361 GAATACTGGTGCCAGTGCCTGGCATGGAGCTCCTGGGCACCAAGAGTCAGAAGGCC 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GAATACTGGTGCCAGTGTGGCATGGAGCTCCTGGGTACCAAGAAAAGTCAGAAGGCC 420

Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACCTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	421	TACATCCGGATTGCCATTGCGCAAGAACCTTGAGCAGGAGCCACTGGCCAAGGAAGTG	480
Qy	481	TCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	481	TCACTGGAGCAAGGCATTGTACTACCTTGTGCCCCCCAGAAGGAATCCCCCAGCTGAG	540
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC	600
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	601	ACGCGGGAGCACAGCCTGGTGGCGACAGGCCCGCCTGCTGACACGGCCAACATACACC	660
Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACATACACC	660
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGGCCCTCCGCTGCTGTCATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCGTGCCGAAGCACCTCTGCAGCGGTATTGTTAT	720
Qy	721	GTGAACGGTGGGTGGTCACGTGGACCGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGC	780
Db	721	GTGAACGGTGGGTGGTCACGTGGACTGAGTGGTCCGTCTGCAGGCCAGCTGTGGCGT	780
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTAACGGGGCGCTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTAACGGGGCGCTTC	840
Qy	841	TGTGAGGGCAGAACATGTCCAGAAAACAGCCTGCGCCACCCCTGTGCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGCAGAACATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCAGTGGATGGAGC	900
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTCGTGGAGTAAGTGGTCAGCCTGTGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGAACGGAGGGAGGAGTGCACGGGACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGAACGGAGGTGAGGAGTGTGGGGTGCTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAAGTGACCTCTGTGTACACAGTGCCTCTGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAAGTGACCTCTGCCCTGCACACCCGCTTGCCCGAGGACGTGGCT	1080
Qy	1081	CTCTATGTGGCCTCATGCCGTGGCGTCTGCCCTGGCTCTGCTGCTGCTTGTCCCTCATC	1140
Db	1081	CTCTACATCGGCCTTGTGCGTGTGGCTGTGCTCTTGTGCTGTTGTGGCCCTTGGGA	1140
Qy	1141	CTCGTTATTGCCGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTACTGTGCAAGAAGGAAGGGCTGGACTCCGATGTGGCGACTCGTCCATCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCGTCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260